2010

Analyses and design strategies for fundamental enabling building blocks: Dynamic comparators, voltage references and on-die temperature sensors

Jun He
Iowa State University

Analyses and design strategies for fundamental enabling building blocks: Dynamic comparators, voltage references and on-die temperature sensors

By

Jun He

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Electrical Engineering

Program of Study Committee:

Degang Chen, Co-major Professor

Randall L.Geiger, Co-major Professor

Robert J. Weber

Chris Chong-Nuen Chu

Kenneth J Koehler

Iowa State University

Ames, Iowa

2010

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Dynamic comparators and voltage references are among the most widely used fundamental building blocks for various types of circuits and systems, such as data converters, PLLs, switching regulators, memories, and CPUs. As thermal constraints quickly emerged as a dominant performance limiter, on-die temperature sensors will be critical to the reliable operation of future integrated circuits. This dissertation investigates characteristics of these three enabling circuits and design strategies for improving their performances.

One of the most critical specifications of a dynamic comparator is its input referred offset voltage, which is pivotal to achieving overall system performance requirements of many mixed-signal circuits and systems. Unlike offset voltages in other circuits such as amplifiers, the offset voltage in a dynamic comparator is extremely challenging to analyze and predict analytically due to its dependence on transient response and due to internal positive feedback and time-varying operating points in the comparator. In this work, a novel balanced method is proposed to facilitate the evaluation of time-varying operating points of transistors in a dynamic comparator. Two types of offsets are studied in the model: (1) static offset voltage caused by mismatches in mobilities, transistor sizes, and threshold voltages, and (2) dynamic offset voltage caused by mismatches in parasitic capacitors or loading capacitors. To validate the proposed method, dynamic comparators in two prevalent topologies are implemented in 0.25 μm and 40 nm CMOS technologies. Agreement between predicted results and simulated results verifies the effectiveness of the proposed method. The new method and the analytical models enable designers to identify the most dominant contributors to offset

and to optimize the dynamic comparators' performances. As an illustrating example, the "Lewis-Gray" dynamic comparator was analyzed using the balanced method and redesigned to minimize its offset voltage. Simulation results show that the offset voltage was easily reduced by 41% while maintaining the same silicon area.

A bandgap voltage reference is one of the core functional blocks in both analog and digital systems. Despite the reported improvements in performance of voltage references, little attention has been focused on theoretical characterizations of non-ideal effects on the value of the output voltage, on the inflection point location and on the curvature of the reference voltage. In this work, a systematic approach is proposed to analytically determine the effects of two non-ideal elements: the temperature dependent gain-determining resistors and the amplifier offset voltage. The effectiveness of the analytical models is validated by comparing analytical results against Spectre simulation results.

Research on on-die temperature sensor design has received rapidly increasing attention since component and power density induced thermal stress has become a critical factor in the reliable operation of integrated circuits. For effective power and thermal management of future multi-core systems, hundreds of sensors with sufficient accuracy, small area and low power are required on a single chip. This work introduces a new family of highly linear on chip temperature sensors. The proposed family of temperature sensors expresses CMOS threshold voltage as an output. The sensor output is independent of power supply voltage and independent of mobility values. It can achieve very high temperature linearity, with maximum nonlinearity around +/- 0.05^{o}C over a temperature range of -20^{o}C to 100^{o}C. A sizing strategy based on combined analytical

analysis and numerical optimization has been presented. Following this method, three circuits A, B and C have been designed in standard 0.18 μm CMOS technology, all achieving excellent linearity as demonstrated by Cadence Spectre simulations. Circuits B and C are the modified versions of circuit A, and have improved performance at the worst corner–low voltage supply and high threshold voltage corner. Finally, a direct temperature-to-digital converter architecture is proposed as a *master-slave hybrid temperature-to-digital converter*. It does not require any traditional constant reference voltage or reference current, it does not attempt to make any node voltage or branch current constant or precisely linear to temperature, yet it generates a digital output code that is very linear with temperature.

CHAPTER 1
INTRODUCTION

For the past four decades, the semiconductor industry has distinguished itself in the rapid pace of improvements in all kinds of products. Since 2005, the concept of "More than Moore" proposed by the International Technology Roadmap for Semiconductors (ITRS) has been discussed. It points out two dimensional trends for semiconductor industry. One dimension is called "More Moore" and contains two main aspects: the traditional geometrical downscaling motivated by Moore's Law, and three-dimensional device structure improvement and novel design technology, such as multi-core design. The other dimension refers to functional diversification or "More than Moore." The "More than Moore" approach stimulates a higher integration of functionalities and provides more value at the customer end. The two trends demonstrate the principal trends in the following categories: integration level, cost, speed, power, compactness, functionality, and reliability [1].

This work focuses on the analyses and designs of three fundamental building blocks—dynamic comparators, voltage references, and on-die temperature sensors. They are widely used in a variety of circuits and systems, such as data converters, phase locked loops (PLLs), memories, switching regulators, and medical sensors. Their performance plays a crucial role in the realization of low power, low cost, high integration, and a good reliable design of the whole system.

ANALYSES OF STATIC AND DYNAMIC OFFSET PREDICTION IN DYNAMIC COMPARATORS

An important portion of the dissertation focuses on the offset analyses of dynamic comparators. As a low power, fast-speed, decision-making circuit, dynamic comparators form the core in flash analog to digital converters (ADCs), play a critical role in pipeline ADCs, and are also applied in a wide range of other analog and mixed-signal circuits and systems. Offset analyses for operational amplifiers (Op-Amps) have been well developed for years and are well documented in textbooks and literatures. However, there is no thorough and accurate analysis on how to evaluate the offset of dynamic comparators. In the existing literature, authors tend to analyze the static input offset voltage in a dynamic comparator in the same way as in a traditional Op Amp [2]–[4]. This approach works well for an op-amp based comparator since the operation regions of all transistors are well defined. However, in dynamic comparators with an internal positive feedback, that method is no longer applicable since the transconductance g_m and output conductance g_o of any transistor in a dynamic comparator are input level dependent and time dependent, and hence, not well defined.

Since the operation region of every transistor in a dynamic comparator is changing during each clock period, and how it changes depends on the input signal value during that clock period, a dynamic comparator is viewed as a time-varying, non-linear system. In this work, a balance mode approach is first proposed to analyze the offset in this time-varying, non-linear system. Following this approach, analytical expressions for offset voltages can be derived once the comparator architecture is given. Two types of mismatches are included in the analytical model: (1) static offset voltages caused by mismatches in μC_{ox}

and the threshold voltage V_{th}; and (2) dynamic offset voltage due to mismatches in parasitic capacitances at various circuit nodes. From the analytical models, designers can obtain an insight about the main contributors to offset and can thus fully explore the tradeoffs in dynamic comparator design, such as offset voltage, area, and speed. To illustrate the potential of the approach, the explicit expressions of offset voltage for static random offset were applied to guide the optimization of the classic "Lewis-Gray" comparator. Compared to the original design, the input offset voltage was easily reduced by 41% after optimization while maintaining the same silicon area.

SYSTEMATIC CHARACTERIZATIONS OF NON-IDEAL EFFECTS IN HIGH PRECISION VOLTAGE REFERENCES

The second topic of this work is the systematic characterizations of non-idealities on the performances of high-precision voltage references. Voltage reference circuits provide an output voltage that is ideally constant and independent of supply variations, process parameter variations, and environmental changes [5]–[9]. Voltage reference circuits have been widely applied in a broad range of systems, such as high performance data converters, PLLs, and monolithic sensors. Given the wide applicability, industry has explored voltage references' characteristics and fabricated them in integrated circuit technology for low cost high volume production.

Bandgap references extract the bandgap voltage from the device characteristics of a pn junction by using either bipolar transistors or diodes. How this extraction is accomplished is strongly dependent upon the circuit in which these junction devices are embedded. As such, designing and analyzing bandgap references requires a good model

of the relationship between the port electrical variables and the temperature for a pn junction.

Despite the fact that performance of reported bandgap voltage references is improving, until recently, a closed-form expression for the output voltage of the most basic bandgap references was not available, making it difficult to determine analytically and systematically the effects of the temperature dependence of non-ideal components on the magnitude of the output voltage, the inflection point location, and the curvature of these bandgap circuits. In this work, several non-ideal components that can adversely affect the performance of bandgap references are identified. A systematic approach is proposed to determine analytically the effects of the temperature dependence of non-ideal components. Analytical expressions for the effects of two of the most common non-ideal components—the temperature-dependent gain-determining resistors and the amplifier offset voltage on the temperature characteristics of basic bandgap circuits—are developed. The effectiveness of the analytical expression is demonstrated by comparing the analytical results with results using Cadence Virtuoso Spectre circuit simulator.

DESIGNS OF HIGHLY LINEAR VERY COMPACT ON-DIE TEMPERATURE SENSORS

The third major part of my work is the design of highly linear, very compact on-die temperature sensors. Driven by higher integration trends in semiconductor roadmaps, the component density per unit die has been dramatically increased, leading to excessive chip heating, causing serious concerns regarding chip reliability and performance [1]. For example, in recent years, the power density in micro-processors and related System on Chip (SoC) scale circuits has been dramatically increasing with decreasing feature sizes

and increasing clock speeds. To alleviate the thermal process, performance improvements in emerging processes are coming in the architectural level using multi-core structures along with power management techniques that include combinations of measurement-driven dynamic supply voltage scaling, dynamic clock frequency scaling, and pre-calculated or dynamic task assignments. The measurement-driven power/thermal management approach is based upon thermal measurements at multiple strategic sites in the cores, as well as at critical locations throughout the remainder of the die. The functions of the on-die temperature sensors are: monitoring the temperature at critical sites, and providing feedbacks from sensory data into techniques for thermal management and system performance optimizations. Although the temperature measurement appears to be a straightforward task, the requirements of the on-die temperature sensor are quite strict. The sensors need to be very small so that they can be placed close to a small group of critical transistors. In addition, the sensor needs to be accurate enough for measuring temperature. For example, if we use the sensor's output as the input to the electric field dependent model (E-model) [11]–[13] to model the time dependent dielectric breakdown (TDDB), it can be readily shown that a 5^{o}C error in temperature measurement at 50^{o}C causes a 34% error in the mean time to failure (MTF) in mature processes. This error could be even worse in processes with thinner oxides. Other failure mechanisms are also adversely affected by temperature measurement errors. Correspondingly, using the widely accepted Black model for electro-migration [14], it can be shown that a 5^{o}C error in temperature measurement at 50^{o}C causes a 34% error in the MTF; a 50% error in current measurement causes a 56% error in the MTF; and a combined temperature current measurement error causes an error of 71% in the MTF. The existing integrated

temperature sensors are not sufficient to fulfill the multi-site high-precision measurement task.

In this work, a family of small temperature sensors is proposed to provide a possible solution for highly accurate, on-die temperature measurements. The circuits are able to extract a complementary metal-to-oxide-semiconductor (CMOS) transistor threshold voltage that is very linear with temperature. A procedure to achieve second and third order temperature nonlinearity compensation has been described. Simulation results demonstrate that a maximum temperature error of around 0.05^{o}C can be achieved, which is more than ten times better than the latest reported CMOS on-die temperature sensor. Finally, a temperature-to-digital sensor cell called the "master-slave hybrid temperature-to-digital converter" is proposed. It does not require any traditional constant reference voltage or reference current, and it does not attempt to make any node voltage or branch current constant or precisely linear to temperature, yet it generates a digital output code that is extremely linear to temperature. All these features make this proposed topology very suitable for on-die, multi-site temperature measurement with high accuracy requirements.

DISSERTATION ORGANIZATION

This dissertation comprises a collection of three published papers, two accepted papers and two chapters that will result in two additional papers. Chapter 1 gives an introduction of analyses and design strategies for three fundamental enabling building blocks: dynamic comparators, voltage references, and on-die temperature sensors. Chapter 2 is based upon a paper published in the *IEEE Transactions on Circuits and Systems*, 2009 and *IEEE International Symposium on Circuits and Systems*, 2008,

presenting the "balanced mode method" to analyze random static offset voltage caused by process variations in threshold voltage V_{th} and ß factor $\mu C_{ox}W/L$ in dynamic comparators. This method facilitates the offset prediction in dynamic comparators—the time-varying, non-linear systems. The proposed analytical model provides guidance in the optimization to reduce offset. Chapter 3 is a paper accepted by the *IEEE International Symposium on Circuits and Systems*, for publication in 2010. It focuses on capacitive mismatch—a significant contributor to overall offset but more challenging to predict analytically previously because of the energy storage feature of the capacitors. In this chapter, the balanced mode method is generalized and applied to estimate capacitive mismatch induced offset. Chapter 4 is a paper published by the *Midwest Symposium on Circuits and Systems*, 2008. It presents a systematic approach to determine the effects of non-ideal components on the performances of high precision voltage references, such as the magnitude of the output voltage, the inflection point location, and the curvature. The Appendix to Chapter 4 is laboratory measurement results of bandgap voltage references using 0.6μm technology on-chip diodes. The measured thermal transfer of this reference circuit demonstrates a positive curvature, opposite to the traditional curvature of the reported bandgap voltage references. Chapter 5 contains a paper published by the *Midwest Symposium on Circuits and Systems* 2010, and it focuses on the analysis and designs of highly linear, very compact, trim-less on-die temperature sensors. Chapter 6 proposes two alternative threshold voltage extraction circuits for temperature sensors with improved immunity to the low-voltage high-threshold corner compared to the structure in Chapter 5. Chapter 7 proposes a new topology of temperature-to-digital converter. This temperature-to-digital converter does not require any conventional

voltage or current references, or contain any traditional analog-to-digital converter. The design provides a possible solution for multi-site, on-die temperature measurements with high accuracy requirements. Chapter 8 presents the overall conclusions of the work.

REFERENCES

[1] Semiconductor Industry Roadmap Executive Summary, "International Technology Roadmap for Semiconductors," 2009.

[2] L. Sumanen, M. Waltari and K. Halonen, "A mismatch insensitive CMOS dynamic comparator for pipeline A/D converters," *Proc. IECES*, pp. I-32-35, Dec. 2000.

[3] L. Sumanen, M. Waltari, V. Hakkarainen and K. Halonen, "CMOS dynamic comparators for pipeline A/D converters," *Proc. ISCAS*, pp.V-157-V160, May 2002.

[4] G. A. Al-Rawi, "A New Offset Measurement And Cancellation Technique For Dynamic Latches," *Proceedings of the 2002 IEEE International Symposium on Circuits and Systems*, Vol. 5, pp. 149-152, May 2002.

[5] D. F. Hilbiber, "A new semiconductor voltage standard," *Dig. Techn. Papers ISSCC*, pp.32-33, Feb. 1964.

[6] R. Widlar, "Designing Positive Voltage Regulators", *EEE*, Vol. 17, pp. 90-97, June 1969.

[7] R. Widlar, "New Developments in IC Voltage Regularors", *IEEE Journal of Solid State Circuits*, Vol. 6, pp. 2-7, Feb. 1971

[8] P. R. Gray and R. G. Meyer, *Analog Integrated Circuits***.** New York: Wiley, 2001.

[9] A. P. Brokaw, "A simple three-terminal IC bandgap reference," *IEEE J. Solid-State Circuits*, Vol. SC-9, pp. 388–393, Dec. 1974.

[10] K. C. Boyko and D. L. Gerlach, “Time Dependent Dielectric Breakdown of 210 Å Oxides,” *International Reliability Physics Symposium*, pp.1-8, April 1989.

[11] J. S. Suehle and P. Chaparala, “Low electric field breakdown of thin SiO2 films under static and dynamic stress,” *IEEE Trans. Electron Devices*, Vol. 44, p. 801, May 1997.

[12] J. Suehle, P. Chaparala, C. Messick, W. Miller, and K. Boyko, “Field and temperature acceleration of time-dependent dielectric breakdown in intrinsic thin SiO_2,” *Proc. IEEE IRPS*, Vol. 32, p. 120-125, 1994.

[13] M. Kimura, “Oxide breakdown mechanism and quantum physical chemistry for time-dependent dielectric breakdown,” *Proc. IEEE IRPS*, 1997, vol. 35, p. 190.

[14] J. R. Black, “Electromigration-A brief survey and some recent results,” *IEEE Trans. Electron Devices*, Vol. ED-16, no. 4, p. 338-347, 1969.

CHAPTER 2
ANALYSES OF STATIC AND DYNAMIC RANDOM OFFSET VOLTAGES IN DYNAMIC COMPARATORS

A paper published in IEEE Transactions on Circuits and Systems

Part I: Regular Papers, Vol.56, No.5, pp.911-919, May 2009

Jun He, Sanyi Zhan, Degang Chen, Randall Geiger

ABSTRACT

When mismatches are present in a dynamic comparator, due to internal positive feedback and transient response, it is always challenging to analytically predict the input referred random offset voltages since the operating points of transistors are time varying. In this paper, a novel balanced method is proposed to facilitate the evaluation of operating points of transistors in a dynamic comparator. Therefore it becomes possible to obtain an explicit expression for offset voltage in dynamic comparators. We include two types of mismatches in the model: (1) static offset voltages from the mismatch in μC_{ox} and threshold voltage V_{th}; (2) dynamic offset voltage due to the mismatch in the parasitic capacitances. From the analytical models, designers can obtain an intuition about the main contributors to offset and also fully explore the tradeoffs in dynamic comparator design, such as offset voltage, area and speed. To validate the balanced method, two topologies of dynamic comparator implemented in 0.25μm and 40nm CMOS technology are applied as examples. Input referred offset voltages are first derived analytically based on SPICE Level 1 model, whose values are compared with more accurate Monte Carlo transient simulations using a sophisticated BSIM3 model. A good agreement between

those two verifies the effectiveness of the balanced method. To illustrate its potential, the explicit expressions of offset voltage were applied to guide the optimization of a "Lewis-Gray" structure. Compared to the original design, the input offset voltage was easily reduced by 41% after the optimization while maintaining the same silicon area.

***Index Terms*—Dynamic Comparators, Monte Carlo Method, Static Offset Voltage, Dynamic Offset Voltage**

I. INTRODUCTION

Comparators have a crucial influence on the overall performance in high-speed analog-to-digital converters (ADCs) [1]. Since they are decision-making circuits that interface the analog and digital signals, the accuracy, which is often determined by its input referred offset voltage, is essential for the resolution of high performance ADCs [2]. Dynamic comparators are widely used in the high speed ADCs due to its low power consumption and fast speed. However, there is a lack of thorough and accurate analysis in the literature on how to evaluate the input offset voltages analytically. Although there exist various offset cancellation circuits and digital calibration techniques [3] [4], to apply such additional circuits to cancel offset voltages increases the power consumption, silicon area and lowers the overall speed. When the transistor feature size is scaled down, random offsets impact the yield of ADCs more severely [5]. Different from the offset caused by mismatch from the gradient effect, random offset cannot be relieved by any layout strategy [6]. In order to achieve an optimum dynamic comparator design, it is essential to have analytical methods to predict offset voltages, especially random offset voltages and provide a deeper insight into the main offset contributors.

Neglecting error sources from external circuits, such as timing error and variation of reference voltages, the offset voltage in a dynamic comparator is mainly comprised of two types of mismatch: (1) static mismatch from variation in μC_{ox} and threshold voltage V_{th}; (2) dynamic mismatch from internal node parasitic capacitors' imbalance. In the previous literatures, both of the mismatches are not well characterized.

First, the previous authors tend to analyze the static input offset voltage in a dynamic comparator in the same way as in the traditional operational amplifier [7]–[9]. The calculation of offset voltage in the op-amp based comparator is straightforward since the operation regions of all transistors are well defined. However, in dynamic comparators with an internal positive feedback, the previous method is not applicable since trans-conductance g_m and output conductance g_o of any transistor are time-dependent and not well defined. The authors fail to clearly state how to determine the value of g_m and g_o of the transistors at time-varying condition.

To overcome the difficulties in determining the operation regions and bias conditions of transistors in a dynamic comparator when the mismatch exists, we previously proposed a balanced method to calculate the static input offset voltage [10]. In this method, we first solve the bias point at the comparison phase when the circuit is perfectly balanced without any mismatch. Then, if any mismatch is involved, we apply a compensation voltage ΔV_{in} at one of the input terminals to cancel the mismatch effect and ensure the comparator to reach the balanced status again. ΔV_{in} is the input referred offset voltage. Its variance σ^2_{Vos} is regarded as the square of random offset voltage. Therefore, analytical expressions for static input offset voltage are derived and allow designers to focus on the most influential offset contributors. In very recent publication about thermal

noise analysis in dynamic comparators [11], authors divided the transient process into three phases and performed noise analysis from stochastic differential equations in each time phase. However, in each phase, it is still not straight-forward to determine the bias points for each transistor. Besides, utilizing the piece-wise linear method takes considerable effort and time.

Second, very little emphasis is placed on mismatch of internal parasitic capacitance. Although the feature size of transistors continues to be scaled down, the associated parasitic capacitance is not necessarily decreased due to the reduction of the oxide thickness and the junction depths [5]. In [12] [13], the authors point out that 1fF or 2fF capacitance mismatch at the output can lead to several tens of milivolts offset. Compared with the absolute capacitance mismatch, our study in this work find that the relative capacitance mismatch defined as $\Delta C/C_n$ is more significant in affecting input offset voltage. ΔC is the capacitor mismatch at differential nodes; C_n is nominal capacitance at those nodes.

The paper is organized as follows. In section II, the balanced method is explained and applied to an example to analyze the static offset voltage from random mismatch in μC_{ox} and V_{th}; then the dynamic offset voltage from internal capacitor mismatch is derived based on the similar approach. In section III, our analytical results are compared with the more accurate Monte Carlo transient simulations. “Lewis-Gray” comparators implemented in 0.25μm and 40nm CMOS technology are built and analyzed. A good agreement between the simulated values and derived values shows the effectiveness of this balanced method. The derived analytical expressions provide deeper insights into the most dominant offset contributors and design tradeoffs. In section IV, the analytical

results are applied to guide the optimization of a dynamic comparator and easily reduce the offset voltage by 41% by re-sizing the transistors while maintaining a constant total area. In section V, the method proposed in section II is further verified by predicting offset voltage for another popular comparator topology built in 40nm CMOS process. Section VI summaries the work.

II. RANDOM OFFSET VOLTAGE IN DYNAMIC COMPARATORS

We use the comparator architecture in Fig.2.1 for our analysis. It is based on the structure reported in [14]. The so called "Lewis-Gray" comparator is a widely used dynamic comparator in pipeline A/D converters. The method we proposed to evaluate offset voltage can be similarly extended to characterize offset in other dynamic comparators. In section V, we applied the proposed method to another popular comparator topology introduced in [7] and developed its analytical model for offset voltage. The simulated results also show a good agreement between the derived results and Monte Carlo simulation results.

Static offset voltage from μC_{ox} and V_{th} mismatches

A fully differential dynamic comparator will maintain a balanced state if no mismatch exists in the circuit. For static offset voltage, balanced state means that $V_{out+}=V_{out-}$; currents I_1 and I_2 in both branches are identical at all the time during the transient process. The balanced state can be described by a space Φ_b comprised of power supplies, external bias voltage V_{latch} and comparison threshold or reference voltages V_{ref+} and V_{ref-} and transistor node voltages, which is written as $\Phi_b = \{V_{DD}, V_{latch}, V_{ref+}, V_{ref-}, V_{s5}$ or V_{s6}, V_{d5} or V_{d6}, V_{out+} or $V_{out-}\}$, in which the subscripts *s* and *d* mean source and

drain voltage of a transistor, respectively. When mismatch occurs, the circuit will lose its balance so that $V_{out+} \neq V_{out-}$. A voltage ΔV_{in} can be applied to compensate the mismatch effect and make V_{out+} equal to V_{out-}. This compensation voltage ΔV_{in} is the input offset voltage. The new balanced state Φ_{bn} is the same as Φ_b, because mismatches are small disturbances that will not change the bias condition of the comparator.

In order to calculate ΔV_{in}, node voltages in the balanced state Φ_b need to be found and then are treated as the desired state when ΔV_{in} is applied to compensate mismatch. The chosen time point to calculate Φ_b is not important since under balanced conditions node voltages for both branches are always symmetrical all the time. In this paper, the time point when the control signal V_{latch} reaches V_{DD} is chosen. Therefore, the operation regions of all of the transistors are well defined. Transistors of M_1-M_4 connecting to the input and reference voltages are in the triode region and act like voltage controlled resistors. M_{10} and M_{11} have equal drain and gate voltage, which makes them work in the saturation region. M_7 and M_8 work as switches embedded in cross-coupled inverter pairs made of M_5M_{10} and M_6M_{11}. They are turned on during comparison phase and are working in the triode region because of their high gate voltage $V_{g7,\ 8} = V_{DD}$. The drain voltage of M_5 and M_6 is pulled up close to V_{out+} or V_{out-}, and M_5 and M_6 work in saturation. M_9 and M_{12} are both turned off because control signal V_{latch} is V_{DD}, which indicates that mismatch effects in M_9 and M_{12} is negligible. Once the operation region for each transistor is known, combining with known power supply voltages, input voltages and process parameters, each node voltage in the dynamic comparator in the balanced state can be readily solved.

If other time point for the analysis is chosen, for instance, when V_{latch} is half of

the V_{DD}, the operation regions of M_7 and M_8 become unclear. In that situation, the operation regions need to be assumed first, and then verified by solving each node voltages under the balanced condition. Iteration may be necessary to find the operation region of M_7 and M_8.

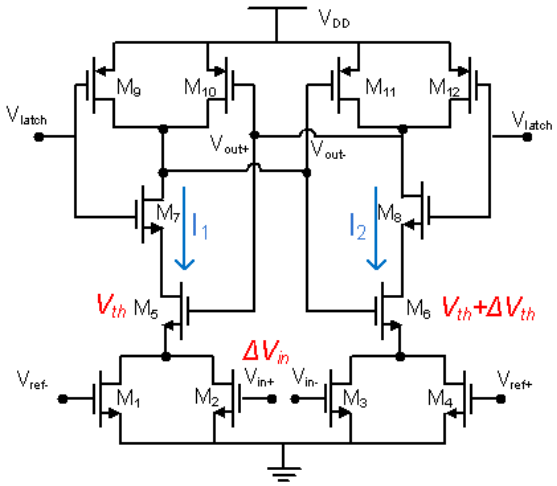

Figure 2.1 "Lewis-Gray" structure

In this paper, mismatch in μC_{ox} and threshold voltage V_{th} are assumed to be the dominant factors in causing the static offset voltage. First, mismatch between M_5 and M_6 is considered and other pairs are assumed to be perfectly matched. Since at the balanced state, by KCL (Kirchhoff's Current Law), the current flowing through M_5 is the sum of currents in M_1 and M_2; the current in M_6 is sum of that in M_3 and M_4. The operation regions of the transistors M_1–M_4 and M_5–M_6 are well defined when the circuit is balanced and analyzed at $V_{latch} = V_{DD}$. M_1–M_4 work as voltage controlled resistors and operate in the triode region. M_5 and M_6 have the drain voltages almost equal to V_{out-} and V_{out+}, respectively, so they are in saturation. Applying a square law model, the currents through M_1–M_4, M_5 and M_6 can be expressed as

$$I_{ds1} = \mu_1 C_{ox} \cdot (\frac{W_1}{L_1}) \cdot (V_{ref-} - V_{t1} - \frac{V_{ds1}}{2}) \cdot V_{ds1} \tag{2.1}$$

$$I_{ds2} = \mu_2 C_{ox} \cdot (\frac{W_2}{L_2}) \cdot (V_{in+} + \Delta V_{in} - V_{t2} - \frac{V_{ds1}}{2}) \cdot V_{ds1} \tag{2.2}$$

$$I_{ds3} = \mu_3 C_{ox} \cdot (\frac{W_3}{L_3}) \cdot (V_{in-} - V_{t3} - \frac{V_{ds3}}{2}) \cdot V_{ds3} \tag{2.3}$$

$$I_{ds4} = \mu_4 C_{ox} \cdot (\frac{W_4}{L_4}) \cdot (V_{ref+} - V_{t4} - \frac{V_{ds3}}{2}) \cdot V_{ds3} \tag{2.4}$$

$$I_{ds5} = \frac{\mu_5 C_{ox}}{2} \cdot \frac{W_5}{L_5} \cdot (V_{out+} - V_{s5} - V_{t5})^2 \tag{2.5}$$

$$I_{ds6} = \frac{\mu_6 C_{ox}}{2} \cdot \frac{W_6}{L_6} \cdot (V_{out-} - V_{s6} - V_{t6})^2 \tag{2.6}$$

When μC_{ox} and threshold voltage V_{th} have mismatch between M_5 and M_6, they can be expressed in terms of a nominal part and a variation part. Since μ and C_{ox} are always in a form of a product, the combined variation can be regarded as the only variation in mobility μ for the convenience of calculation.

$$\mu_5 = \mu_n + \Delta\mu_5 \tag{2.7}$$

$$\mu_6 = \mu_n + \Delta\mu_6 \tag{2.8}$$

$$V_{t5} = V_{th} + \Delta V_{t5} \tag{2.9}$$

$$V_{t6} = V_{th} + \Delta V_{t6} \tag{2.10}$$

where μ_n and V_{th} are the nominal values of NMOS mobility and threshold voltage, respectively. $\Delta\mu_5$ and $\Delta\mu_6$ are the mobility variations for M_5 and M_6. ΔV_{t5} and ΔV_{t6} are

the variations in threshold voltages of M_5 and M_6, respectively.

As a compensation voltage to ensure the comparator work in the balanced condition Φ_b, ΔV_{in} is the offset voltage V_{os_M5M6} caused by mismatch between M_5 and M_6. It can be written as function of mobility μ and threshold voltage V_{th} based on (2.1)-(2.10),

$$V_{os\_M5M6} = \frac{(\mu_n + \Delta\mu_5)(W_5/W_1)}{2\mu_n V_{s5}}(V_{out+} - V_{s5} - V_{tn} - \Delta V_{t5})^2 \; - \frac{(\mu_n + \Delta\mu_6)(W_6/W_1)}{2\mu_n V_{s6}}(V_{out-} - V_{s6} - V_{tn} - \Delta V_{t6})^2 \tag{2.11}$$

where V_{out+}, V_{out-}, V_{s5}, V_{s6} are solved node voltages in the balanced state Φ_b and $V_{out+}=V_{out-}$, $V_{s5}=V_{s6}$. For matching purpose, all the transistors are sized to have the same channel length L; and the four input transistors M_1–M_4 are sized in the same dimensions.

In a practical application, the variation part ΔV_t and $\Delta\mu$ is normally very small compared to the nominal component if reasonable yield is to be guaranteed [6]. Therefore, it can be derived that offset voltage from M_5M_6 mismatch can be approximated in the expression

$$V_{os\_M5M6} \approx K_{VT5}\cdot\Delta V_{t5} + K_{VT6}\cdot\Delta V_{t6} + K_{\mu5}\cdot\Delta\mu_5/\mu_n + K_{\mu6}\cdot\Delta\mu_6/\mu_n \tag{2.12}$$

where $K_{VT5} = -\frac{V_{out+} - V_{s5} - V_{th}}{V_{s5}}\cdot\frac{W_5}{W_1}$ (2.13)

$$K_{VT6} = \frac{V_{out-} - V_{s6} - V_{th}}{V_{s6}}\cdot\frac{W_6}{W_1} \tag{2.14}$$

$$K_{\mu5} = \frac{(V_{out+} - V_{s5} - V_{th})^2}{2\cdot V_{s5}}\cdot\frac{W_5}{W_1} \tag{2.15}$$

$$K_{\mu6} = -\frac{(V_{out-} - V_{s6} - V_{th})^2}{2\cdot V_{s6}}\cdot\frac{W_6}{W_1} \tag{2.16}$$

In the BSIM3 and BSIM4 models, mobility μ and threshold voltage V_{th} have a weak correlation in high order terms [15][16]. To simplify the derivation, we assume that μ and V_{th} are uncorrelated with each other and have a nearly Gaussian distribution. It is well known that the linear combination of Gaussian random variables is Gaussian [6]. Random offset voltage caused by mismatch from M_5 and M_6 can be derived from the variance of (2.12)

$$\sigma^2_{V_{OS\_M5M6}} = K^2_{VT5}\cdot\sigma^2_{V_{t5}} + K^2_{VT6}\cdot\sigma^2_{V_{t6}} + K^2_{\mu_5}\cdot\sigma^2_{\mu_5/\mu_n} + K^2_{\mu_6}\cdot\sigma^2_{\mu_6/\mu_n} \tag{2.17}$$

where K^2_{VT5}, K^2_{VT6}, $K^2_{\mu 5}$ and $K^2_{\mu 6}$ have been expressed in (2.13)-(2.16). Similarly, input random offset voltages caused by mismatch of the other pairs can also be found as follows. Random offset from mismatch between M_1 and M_4 is

$$\sigma^2_{V_{os\_M1M4}} = \sigma^2_{V_{t1}} + \sigma^2_{V_{t4}} + (V_{ref+} - V_{tn})^2\cdot\sigma^2_{V_{\mu_4/\mu_n}} + (V_{ref-} - V_{tn})^2\cdot\sigma^2_{V_{\mu_1/\mu_n}} \tag{2.18}$$

Random offset from mismatch between M_2 and M_3 is

$$\sigma^2_{V_{os\_M2M3}} = \sigma^2_{V_{t2}} + \sigma^2_{V_{t3}} + (V_{in+} - V_{tn})^2\cdot\sigma^2_{V_{\mu_2/\mu_n}} + (V_{in-} - V_{tn})^2\cdot\sigma^2_{V_{\mu_3/\mu_n}} \tag{2.19}$$

Random offset from mismatch between M_{10} and M_{11} is

$$\sigma^2_{V_{os\_M10M11}} = \frac{(V_{DD} - V_{out+} - V_{tp})^2}{9V^2_{s5}}\sigma^2_{V_{t10}} + \frac{(V_{DD} - V_{out+} - V_{tp})^2}{9V^2_{s6}}\cdot\sigma^2_{V_{t11}} + \frac{(V_{DD} - V_{out+} - V_{tp})^4}{36\cdot V^2_{s5}}\sigma^2_{\mu_{10}/\mu_p} + \frac{(V_{DD} - V_{out+} - V_{tp})^4}{36\cdot V^2_{s6}}\sigma^2_{\mu_{11}/\mu_p} \tag{2.20}$$

Random offset from mismatch between M_7 and M_8 is

$$\sigma^2_{V_{os\_M7M8}} = (\frac{W_7}{W_2}\cdot\frac{V_{ds7}}{V_{s5}})^2\cdot\sigma^2_{V_{t7}} + (\frac{W_8}{W_2}\cdot\frac{V_{ds8}}{V_{s5}})^2\cdot\sigma^2_{V_{t8}} + (\frac{W_7}{W_2}\cdot\frac{V_{ds7}}{V_{s5}})^2\cdot(V_{latch} - V_{d5} - V_{tn} - \frac{V_{ds7}}{2})^2\cdot\sigma^2_{V_{\mu 7}/\mu_n}$$

$$+ (\frac{W_8}{W_2}\cdot\frac{V_{ds8}}{V_{s5}})^2\cdot(V_{latch} - V_{d6} - V_{tn} - \frac{V_{ds8}}{2})^2\cdot\sigma^2_{V_{\mu 8}/\mu_n} \qquad (2.21)$$

where V_{tp} is the nominal value for the threshold voltage of PMOS; σ^2_{Vti}, $\sigma^2_{\mu i/\mu n}$ and $\sigma^2_{\mu i/\mu p}$ (i=1,2..11) characterize random mismatch in threshold voltage and mobility of NMOS and PMOS transistors, which can be modeled as follows [5][17],

$$\sigma^2_{V_t} = \frac{A^2_{V_t}}{W\cdot L} + S^2_{V_{T0}}\cdot D^2 \qquad (2.22)$$

$$\sigma^2_{\mu_i/\mu_n} = \frac{A^2_{\mu}}{W\cdot L} + S^2_{\mu}\cdot D^2 \qquad (2.23)$$

where W and L are the width and length of transistors in the pairs. A_{vt}, A_β, S^2_{VT0} and S^2_μ are process dependent parameters and S^2_{VT0}, S^2_μ describe the variation of V_{T0} and μ with the spacing. D is the distance on chip between the matching transistors, which will be neglected because of its minor contribution to the overall mismatch.

If the random mismatches in each pair are uncorrelated or nearly uncorrelated, the overall static random offset voltage σ_{Vos} from mismatch in μC_{ox} and threshold voltage V_{th} in the dynamic comparator can be described as follows:

$$\sigma^2_{V_{os}} = \sigma^2_{V_{os\_M5M6}} + \sigma^2_{V_{os\_M1M4}} + \sigma^2_{V_{os\_M2M3}} + \sigma^2_{V_{os\_M7M8}} + \sigma^2_{V_{os\_M10M11}} \qquad (2.24)$$

Static random offset resulting from mismatch between M_9 and M_{12} is neglected in the calculation, because they work as switches during the reset state to pull up the differential output to V_{DD}, and then are turned off during the comparison stage.

From offset expressions (2.17) to (2.21), we can have the following conclusions about the comparator in Fig.2.1:

--Static random offset voltages caused by mismatch in transistors pairs of M_1 and M_4, M_2 and M_3 can be reduced by increasing the size of those transistors, because σ^2_{Vti} and $\sigma^2_{\mu i/\mu n}$ (i=1,2,3,4) are inversely proportional to the product of W and L;

--Random offset voltages caused by mismatch in transistors pairs of M_5 and M_6, M_7 and M_8 can not be guaranteed to be reduced when the sizes of the transistors are increased since the widths also appear in the numerator of the (2.13)-(2.16) and (2.21).

--A particular aspect ratio W/L can be found to make an optimum tradeoff between random offset voltage and transistor size denoted by the product of W and L, which is discussed in detail in section IV.

Dynamic offset voltage from internal capacitor mismatch

Distinguished from mismatch caused by μC_{ox} and threshold voltage V_{th}, the effects of parasitic capacitance mismatch are shown only during the transient process and therefore are called the dynamic offset. A four-terminal MOS device includes twelve different parasitic capacitors [16]. For a matched pair in the dynamic comparator, any dimension mismatch due to process variation and asymmetric interconnection will cause capacitance mismatch. It has been demonstrated that a 1fF capacitance mismatch at the output node may contribute several tens of mili-volts input-referred offset voltage [13]. For a simple two-inverter latch structure, the authors in [13] have shown analytically the derivation of input referred offset voltage. For the more complicated dynamic comparator as shown in Fig.2.2, an accurate analysis like what they proposed will be very tedious.

As shown in Fig.2.2, C_5 and C_6 contain all of the parasitic capacitance from V_{out-} and V_{out+} to ground, respectively. An accurate analysis to calculate the offset voltage due to capacitor mismatch in the dynamic comparator is tedious since we have to consider the transient current and voltage due to capacitance charge and discharge.

By using the balanced method, a simple formula to calculate the input referred offset voltage due to C_5 and C_6 mismatch using a square law model is derived as follows. In order to calculate the offset voltage, a DC voltage $V_{OS,C56}$ is virtually added to the V_{in+} terminal voltage. When the comparator is balanced, V_{out+} and V_{out-} are equal and dV_{out+}/dt and dV_{out-}/dt are equal. The time point to calculate the dynamic offset voltage is chosen at when M_{10} and M_{11} are about to be turned on.

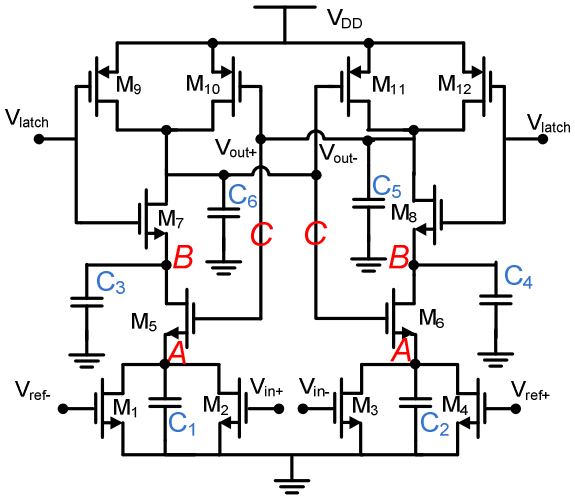

Figure 2.2 "Lewis-Gray" structure with internal parasitic capacitors

Therefore, the compensation DC voltage $V_{OS,C56}$ is the input-referred dynamic offset voltage. To make the formula more readable, we assume that the transient currents through the parasitic capacitance except C_5 and C_6 are negligible. Then the following equations can be written,

$$C_5 \frac{dV_{out-}}{dt} = I_{ds1} + I_{ds2} \tag{2.25}$$

$$C_6 \frac{dV_{out+}}{dt} = I_{ds3} + I_{ds4} \tag{2.26}$$

$$\frac{dV_{out+}}{dt} = \frac{dV_{out-}}{dt} \tag{2.27}$$

From (2.25) ~ (2.27), it follows that

$$\frac{I_{ds1} + I_{ds2}}{C_5} = \frac{I_{ds3} + I_{ds4}}{C_6} \tag{2.28}$$

Applying a square law model to replace the drain source current in (2.28), the input referred dynamic offset $V_{os,C56}$ due to mismatch in C_5 and C_6 is derived as,

$$V_{os,C56} = \frac{\Delta C_{56}}{C_6} \frac{I_{ds3} + I_{ds4}}{\mu_n C_{ox} \frac{W_2}{L_2} V_{ds1}} \tag{2.29}$$

where $\Delta C_{56} = C_5 - C_6$

From (2.29), it is shown that the dynamic offset voltage is more affected by the relative capacitance mismatch $\Delta C_{56}/C_6$ than just the absolute capacitance mismatch ΔC_{56}. If the relative capacitance mismatch is decreased at output nodes, the input referred offset voltage will be reduced. A possible strategy to minimize the dynamic offset voltage is increasing the transistor area of the M_5M_6 pair so that the relative mismatch is reduced. Moreover, if the comparator speed requirement can be easily met, some precision capacitors with very good matching properties can be added at the output nodes to further

shrink the relative capacitor mismatch. The Monte Carlo simulations in section III confirm the above conclusions.

III. NUMERICAL EXAMPLES AND MONTE CARLO SIMULATION RESULTS

The previous analysis is validated with simulations results in this section. The "Lewis-Gray" comparator is implemented in both the 0.25μm and 40nm CMOS processes. The key values are listed in Table 2.1&2.1. For better matching purpose, transistor length is chosen to be the same within each process.

Simulations results for static offset voltage

First, all node voltages are solved when no mismatch is present. The bias conditions in the balanced state Φ_b can be determined. In 0.25μm comparator, it can be calculated that: $V_{out+}=V_{out-}=0.601V$, $V_{d5}=V_{d6}=0.585V$, $V_{s5}=V_{s6}=0.0089V$. In the 40nm comparator, the bias condition is calculated as: $V_{out+}=V_{out-}=0.458V$, $V_{d5}=V_{d6}=0.439V$, $V_{s5}=V_{s6}=0.0129V$. Then the calculated node voltages are applied to (2.17)–(2.21) to find numerical values for random offset caused by mismatch due to process variation in each pair. A_{vt} and A_μ in σ_{Vt} and $\sigma_{\mu i/\mu n}$ are process dependent parameters, whose values for different processes are listed as a reference in Table 2.3 [5].

Monte Carlo transient simulation is performed using a BSIM3 model. In the model file, the mobility μ_n and threshold voltage V_{th} are defined as Gaussian distributed variables with a standard deviation modeled by (2.22) and (2.23). One hundred iterations are done for each pair while assuming no mismatch exists in other pairs so that σ_{Vos_M1M4}, σ_{Vos_M2M3}, σ_{Vos_M5M6}, σ_{Vos_M7M8}, σ_{Vos_M10M11} and σ_{Vos_M9M10} can be determined one by one. In Fig.2.3 and Fig.2.4, the random offset voltage calculated by the analytical method shows a good agreement with the Monte Carlo simulation results. From the plot, we can easily tell the most influential offset contributors.

Table 2.1 Key values for the dynamic comparator in 0.25µm

Process	**0.25µm CMOS**
Power supply	V_{DD}=1.5V, V_{ss}=0V
Transistor sizing	$(W/L)_{1,2,3,4}$=(1.5µ/0.45µ)x4
	$(W/L)_{5,6,7,8}$=(3.5µ/0.45µ)x4
	$(W/L)_{10,11}$=(1.5µ/0.45µ)x4
V_{ref}	V_{ref+}=1.6V, V_{ref-}=1.2V
Clock signal V_{latch}	High=1.5V;Low=0V
	Rise and fall time = 10ps
	Pulse width=20ns; Freq=10MHz
Switch (PMOS)	$(W/L)_{9,12}$=1.5µ/0.45µ

Table 2.2 Key values for the dynamic comparator in 40nm

Process	**40nm CMOS**
Power supply	V_{DD}=1.0V, V_{ss}=0V
Transistor sizing	$(W/L)_{1,2,3,4}$=(0.16µ/0.05µ)x4
	$(W/L)_{5,6,7,8}$=(0.38µ/0.05µ)x4
	$(W/L)_{10,11}$=(0.16µ/0.05µ)x4
V_{ref}	V_{ref+}=0.8V, V_{ref-}=0.6V
Clock signal V_{latch}	High=1.0V;Low=0V
	Rise and fall time = 1ps
	Pulse width=500ps; Freq=1GHz
Switch (PMOS)	$(W/L)_{9,12}$=0.16µ/0.05µ

Table 2.3 Mismatch parameters for several CMOS technologies

Technology	Type	A_{vt}(mV·μm)	A_{β}(%·μm)
2.5 μm	NMOS	30	2.3
	PMOS	35	3.2
1.2 μm	NMOS	21	1.8
	PMOS	25	4.2
0.7 μm	NMOS	13	1.9
	PMOS	22	2.8
0.5 μm	NMOS	11	1.8
	PMOS	13	2.3
0.35 μm	NMOS	9	1.9
	PMOS	9	2.25
0.25 μm	NMOS	6	1.85
	PMOS	6	1.85
40 nm	NMOS	1.8	0.45
	PMOS	1.7	0.68

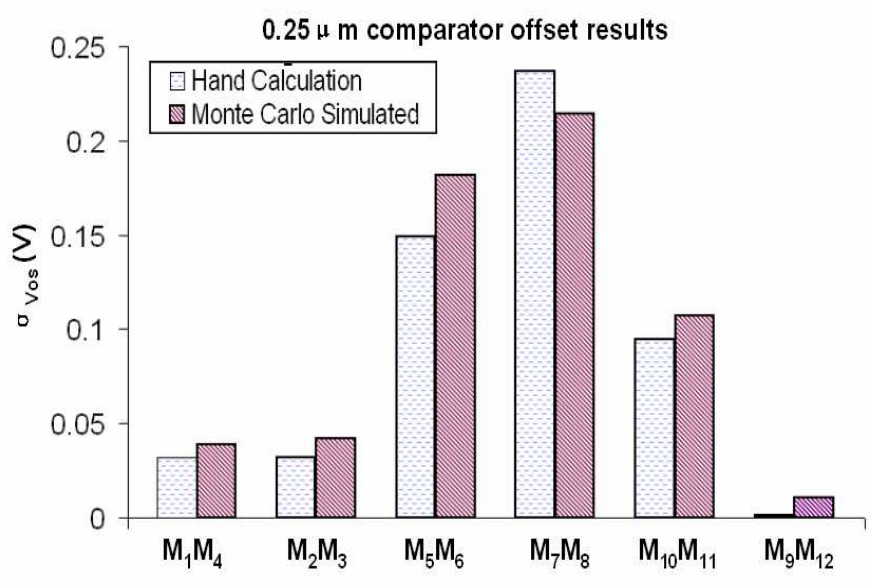

Figure 2.3 Comparison between analytical results and Monte Carlo simulation for each pair in 0.25µm comparator

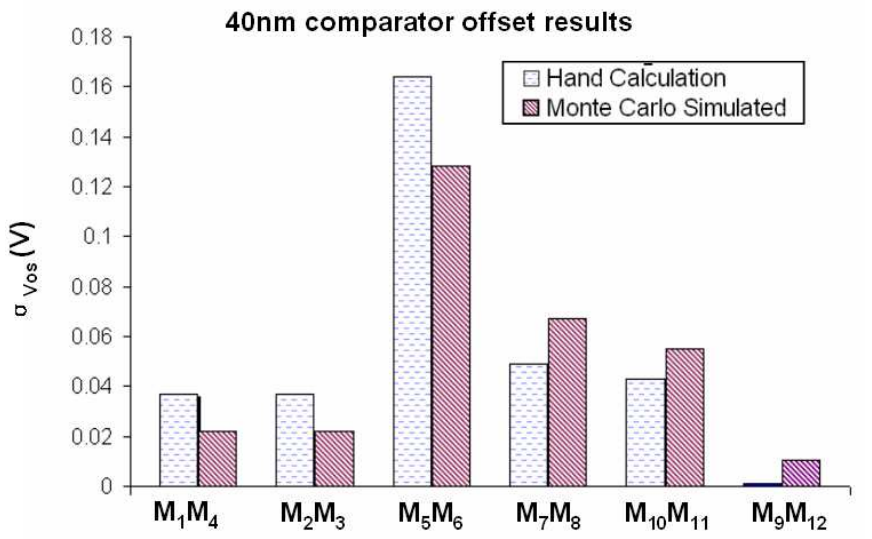

Figure 2.4 Comparison between analytical results and Monte Carlo simulation for each pair in 40nm comparator

Simulations results for dynamic offset voltage

In section II, the explicit expression of dynamic offset voltage due to capacitance mismatch at the output nodes has been derived in (2.29). To demonstrate its effectiveness, Monte Carlo transient simulations are conducted based upon the comparator described in Table 2.1. As we have predicted from the analytical model, it is the relative capacitance mismatch $\Delta C/C_n$ rather than just the absolute mismatch ΔC that plays a key role in determining the dynamic offset voltage. As shown in Fig.2.5, it is clear that as the output nominal capacitance is increasing but the absolute mismatch capacitance is kept as a constant 0.1fF, the dynamic offset is decreasing. The calculated and simulated dynamic offset voltages show a reasonable agreement.

For the dynamic comparator illustrated in Fig.2.2, the contributions to offset voltage from parasitic capacitors at different nodes are usually different. In order to compare the effects of the capacitance mismatch at different nodes, the mismatch capacitor ΔC is added to one of the three nodes *A*, *B* and *C* and the relative capacitance mismatch is kept to be 1.67%. The Monte Carlo simulation demonstrates that capacitance mismatch at the output node *C* accounts for 76% overall dynamic offset voltage. The capacitance mismatch of $C_{1,2}$ and $C_{3,4}$ contributes the remaining dynamic offset voltage. It shows that capacitor mismatch at the output node is the most influential contributor to the dynamic offset.

To probe more, we further investigate the capacitor mismatch at differential nodes A and B. As reported in Fig.2.6, when the nominal capacitance at differential node A, B increases, the dynamic offset voltages increase. The results are predictable since a larger capacitor will dump larger transient currents to ground. The transient currents flowing

through input pairs $M_{2,3}$ are reduced as a result. To cancel the effects of capacitance mismatch in differential nodes, a larger input voltage ΔV_{in} is required to compensate for the mismatch. It suggests that the internal capacitance at nodes A and B should be kept as small as possible.

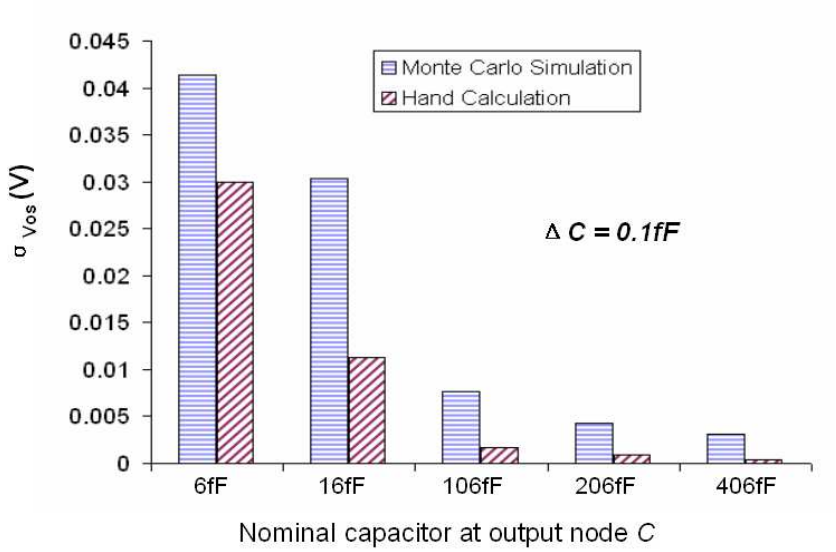

Figure 2.5 Comparison of dynamic offset voltages due to $C_{5,6}$ mismatch derived by analytical model and Monte Carlo simulation

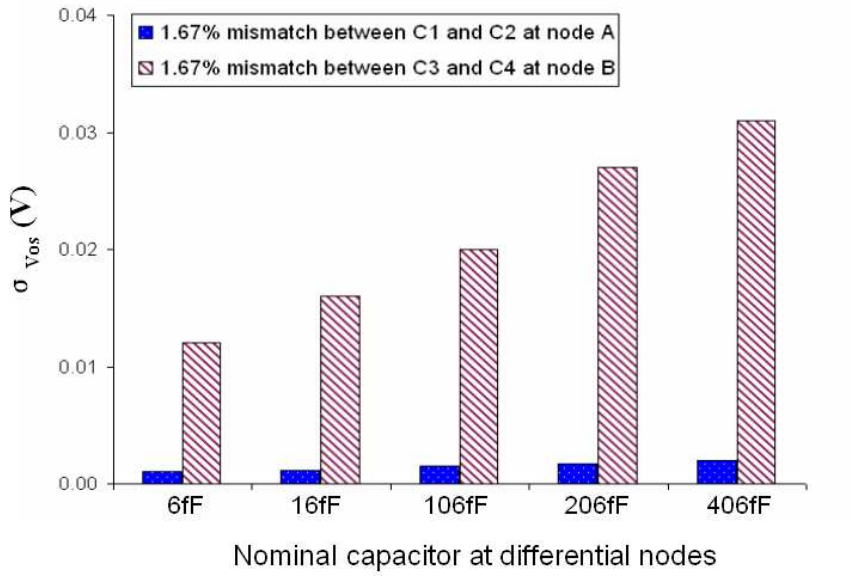

Figure 2.6 Dynamic offsets from C_1&C_2 and C_3&C_4 mismatch

IV. ONE APPLICATION OF THE RANDOM OFFSET ANALYTICAL MODEL

Without any offset cancellation technique, a dynamic comparator will not easily achieve input offset voltage less than several tens of milli-volts. Mismatch caused by random variations cannot be relieved from any layout strategy. From pervious sections, it is indicted that by using a symmetric layout, well-balanced routing and extra balanced capacitive loading, the dynamic mismatch at sensitive nodes—output nodes can be reduced. By contrast, it seems more difficult for designers to control the random mismatch from μC_{ox} and V_{th}. In fact, by utilizing the analytical model in the analytical model in (2.17)–(2.21), the static random offset voltage can be reduced by proper sizing without increasing the total area of the comparator.

The following procedures are applied to find the proper sizes to achieve smaller random offset voltage given a fixed total area.

1) Based on the analytical results in (2.17)-(2.21), the input random offset voltage due to each transistor pair can be calculated. Then all the transistor pairs are divided into several groups following the rule that in each group there contains both a critical matching pair and uncritical pairs.

2) First focus on the mismatch in one group and assume there is no mismatch in the other groups. Based on the conclusion from section II, a minimum random offset voltage can be found by properly adjusting the sizes of the transistor pairs depending on their contributions to the offset voltages. Apply the same procedure to the remaining groups to achieve minimum random offset in each group.

Apply the above procedure to the comparator example described in Table 2.1. Based on the calculated offset voltage from each transistor pair, six pairs of transistors

are divided into two groups. Group 1 is composed of the bottom four uncritical matching transistor pairs M_1–M_4 and critical matching transistor pairs M_7M_8. Group 2 includes the four uncritical matching PMOS transistors M_9–M_{12} and critical matching NMOS pairs M_5M_6.

First, group 1 is optimized. The area budget is moved from M_1–M_4 to M_7M_8 by increasing the width of M_7M_8 in a step size of 0.5μm while the total area in the group is maintained as a constant. The simulated random offset voltage versus width change in M_7M_8 denoted by $\Delta W\_M_7M_8$ is shown in Fig.2.7. It is shown that when $\Delta W\_ M_7M_8$ is equal to 2μm, which means the widths of M_7M_8 are increased from 3.5μm to 5.5μm and widths of M_1–M_4 are decreased from 1.5μm to 0.5μm. The random offset voltage reaches the minimum value 78.3mV within group 1. A similar area allocation procedure is applied to group 2 made up of M_5M_6 and M_9–M_{12}. The simulated random offset versus ΔW of M_5M_6 is shown in Fig.2.8. Finally, the sizes are optimized and listed in Table 2.4. After this optimization, Monte Carlo simulation is applied with mismatch presented in all the pairs, and the overall random offset voltage is 150 mV, which is reduced by 41% compared with 254 mV in the original sizing. The total area is still kept as a constant.

V. MODEL VALIDATION IN COMPARATOR TOPOLOGY II

To further validate the effectiveness of our method in section II, in this section we present another dynamic comparator topology in Fig. 2.9 and apply the method to analyze its offset. The topology is first introduced in [7].

The operation of the comparator can simply be described as follows. When latch signal reaches zero, M_5 and M_6 are turned off and current paths are cut off. M_9 and M_{12} reset the differential output to V_{DD}. When latch signal is raised high, differential output

nodes are disconnected from V_{DD}. Depending on the difference between the input voltage and reference voltage, cross coupled inverter pairs made up of M_7M_{10} and M_8M_{11} regeneratively amplify the difference and determine which of the outputs goes to V_{DD} and which to 0V.

As detailed in section II, to find out the offset voltage from mismatch in μC_{ox} and threshold voltage V_{th}, we will first determine the bias conditions at perfectly balanced condition. We choose $V_{latch}=V_{DD}$ as the time point for analysis. M_7, M_8, M_{10} and M_{11} all have the same gate and drain voltages since $V_{out+}=V_{out-}$ in the balanced state. Therefore, they are working in the saturation region. M_5 and M_6 work as tail current sources. They are supposed to be working in saturation to eliminate large offset due to the mismatch [8]. To avoid M_5 and M_6 goes into triode when latch signal goes high, instead of using the clock going from 0 to V_{DD}, a lower voltage clock V_{latch2} is used to guarantee that M_5 and M_6 remain in saturation [12]. M_1–M_4 are in triode region and act like voltage controlled resistors.

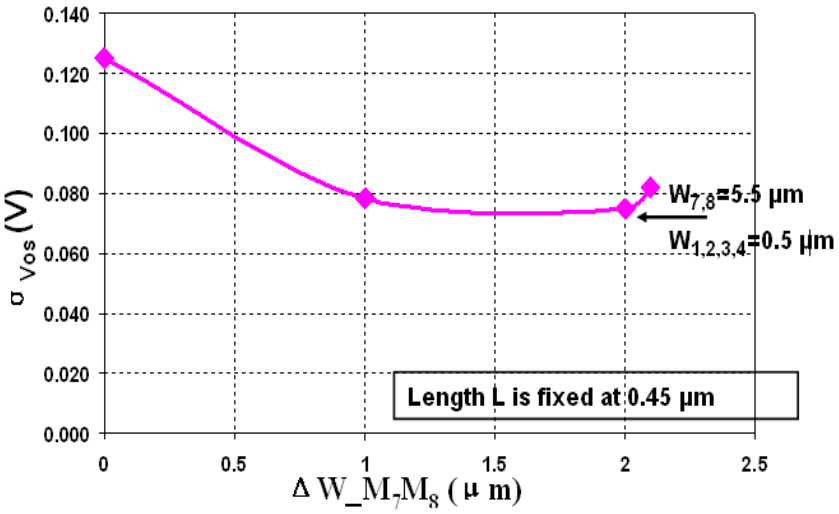

Figure 2.7 Random offset vs. ΔW of matching critical pair M_7M_8

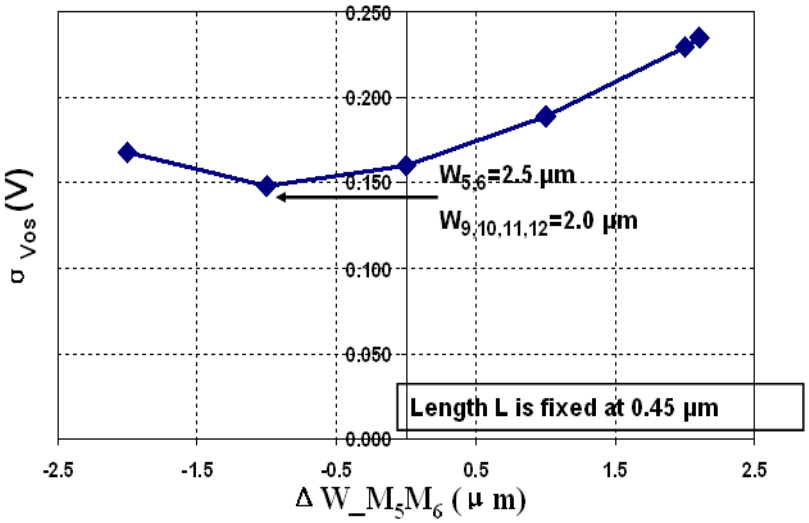

Figure 2.8 Random offset vs. ΔW of matching critical pair M_5M_6

Table 2.4 Dimensions and random offset comparison

Name	Original sizes (µm/ µm)	Optimized sizes (µm/ µm)
M_1M_3	1.5/0.45	0.5/0.45
M_2M_4	1.5/0.45	0.5/0.45
M_5M_6	3.5/0.45	2.5/0.45
M_7M_8	3.5/0.45	5.5/0.45
$M_{10}M_{11}$	1.5/0.45	2.0/0.45
M_9M_{12}	1.5/0.45	2.0/0.45
σ_{Vos}	*254 mV*	*150 mV*

Once the operation regions are determined, by KCL, we can calculate the offset voltage caused by each pair in the comparator. Random offset from mismatch between M_2 and

M_3 is

$$\sigma^2_{V_{os\_M2M3}} = \sigma^2_{V_{t2}} + \sigma^2_{V_{t3}} + (V_{in-} - V_{d5} - V_{tn} - \frac{V_{ds2}}{2})^2 \cdot \sigma^2_{V_{\mu_2/\mu_n}} + (V_{ref-} - V_{d6} - V_{tn} - \frac{V_{ds3}}{2})^2 \cdot \sigma^2_{V_{\mu 3/\mu_n}} \quad (2.30)$$

Random offset from mismatch between M_1 and M_4 is

$$\sigma^2_{V_{os\_M1M4}} = \sigma^2_{V_{t1}} + \sigma^2_{V_{t4}} + (V_{in+} - V_{d5} - V_{tn} - \frac{V_{ds1}}{2})^2 \cdot \sigma^2_{V_{\mu 1/\mu_n}} + (V_{ref+} - V_{d6} - V_{tn} - \frac{V_{ds4}}{2})^2 \cdot \sigma^2_{V_{\mu 4/\mu_n}} \quad (2.31)$$

Random offset from mismatch between M_5 and M_6 is

$$\sigma^2_{V_{os\_M5M6}} = (\frac{W_5}{W_1})^2 \frac{(V_{latch2} - V_{tn})^2}{V^2_{ds1}} \sigma^2_{V_{t5}} + (\frac{W_6}{W_1})^2 \frac{(V_{latch2} - V_{tn})^2}{V^2_{ds1}} \sigma^2_{V_{t6}} + (\frac{W_5}{W_1})^2 \frac{(V_{latch2} - V_{tn})^4}{4 \cdot V^2_{ds1}} \sigma^2_{\mu_5/\mu_n}$$

$$+ (\frac{W_6}{W_1})^2 \frac{(V_{latch2} - V_{tn})^4}{4 \cdot V^2_{ds1}} \sigma^2_{\mu_6/\mu_n} \quad (2.32)$$

Random offset from mismatch between M_7 and M_8 is

$$\sigma^2_{V_{os\_M7M8}} = (\frac{W_8}{W_1})^2 \frac{(V_{out} - V_{s8} - V_{tn})^2}{V^2_{ds1}} \sigma^2_{V_{t8}} + (\frac{W_7}{W_1})^2 \frac{(V_{out+} - V_{s7} - V_{tn})^2}{V^2_{ds1}} \sigma^2_{V_{t7}}$$

$$+ (\frac{W_8}{W_1})^2 \frac{(V_{out-} - V_{s8} - V_{tn})^4}{4 \cdot V^2_{ds1}} \sigma^2_{\mu_8/\mu_n} + (\frac{W_7}{W_1})^2 \frac{(V_{out+} - V_{s7} - V_{tn})^4}{4 \cdot V^2_{ds1}} \sigma^2_{\mu_7/\mu_n} \quad (2.33)$$

Random offset from mismatch between M_{10} and M_{11} is

$$\sigma^2_{V_{os\_M10M11}} = (\frac{W_{10}}{W_1})^2 \frac{(V_{DD} - V_{out+} - V_{tn})^2}{4 \cdot V^2_{ds1}} \sigma^2_{V_{t10}} + (\frac{W_{11}}{W_1})^2 \frac{(V_{DD} - V_{out-} - V_{tn})^2}{4 \cdot V^2_{ds1}} \sigma^2_{V_{t11}}$$

$$+ (\frac{W_{10}}{W_1})^2 \frac{(V_{DD} - V_{out+} - V_{tn})^4}{16 \cdot V^2_{ds1}} \sigma^2_{\mu_{10}/\mu_n} + (\frac{W_{11}}{W_1})^2 \frac{(V_{DD} - V_{out-} - V_{tn})^4}{16 \cdot V^2_{ds1}} \sigma^2_{\mu_{11}/\mu_n} \quad (2.34)$$

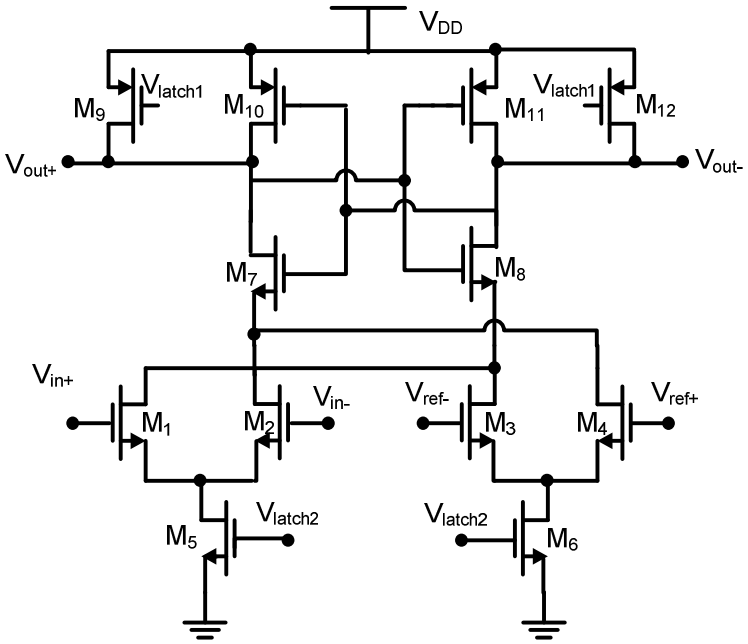

Figure 2.9 Topology II dynamic comparator

Table 2.5 Key values for comparator II in 40nm

Process	**40nm CMOS**
Power supply	V_{DD}=1.0V, V_{ss}=0V
Transistor sizing	$(W/L)_{1,2,3,4,5,6}$=(0.16µ/0.05µ)x4
	$(W/L)_{7,8}$=(0.38µ/0.05µ)x4
	$(W/L)_{10,11}$=(0.16µ/0.05µ)x4
V_{ref}	V_{ref+}=0.8V, V_{ref-}=0.6V
Clock signal V_{latch}	V_{latch1} High=1.0V;Low=0V V_{latch2} High=0.4V; Low=0V
	Rise and fall time = 1ps
	Pulse width=500ps; Freq=1GHz
Switch (PMOS)	$(W/L)_{9,12}$=0.16µ/0.05µ

The topology II dynamic comparator implemented in 40nm operates at a 1.0 GHz clock frequency with a 1.0 V power supply. Table V shows key design parameters. The bias condition at each node is solved as: V_{out+}=V_{out-}=0.579V, V_{s7}=V_{s8}=0.193V, V_{d5}=V_{d6}=0.179V. The calculated node voltages are then applied to (2.30)–(2.34) to find numerical value for random offset caused by mismatch due to process variation in each pair. The calculated values are plotted with the Monte Carlo transient simulation results as a comparison. It can be seen that the analytical result gives a good prediction in the offset voltage from each pair and especially in the main offset contributors.

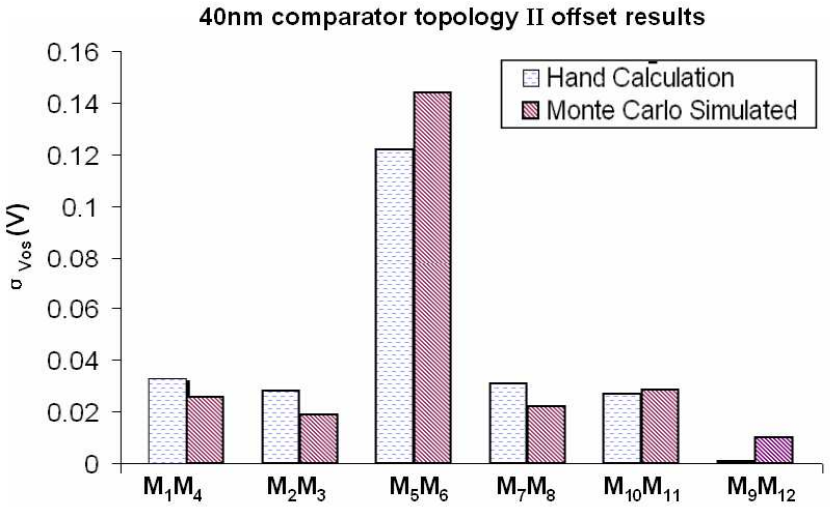

Figure 2.10 Comparison between analytical results and Monte Carlo simulation results for each pair in 40nm comparator in Fig.2.9

VI. CONCLUSIONS

In this paper, we presented a novel balanced method to analyze input referred offset voltages in dynamic comparators. The method solves the problem that in a dynamic comparator the operating points of transistors are not well defined in the transient process. Based on this method, the explicit expressions for static offset voltages

caused by μC_{ox} and V_{th} variation and dynamic offset voltages caused by capacitance mismatch are derived based upon "Lewis-Gray" dynamic comparator. The comparator is implemented in both 0.25μm and 40nm CMOS process as examples. The analytical results from those expressions achieve good agreements with more accurate Monte Carlo transient simulations. The analytical model also gives a good prediction to the offset in the second topology dynamic comparator. Those explicit formulas of offset voltages allow designers to find out the most dominant contributors to offset and to use those formulas as guidance to design and optimize dynamic comparators.

REFERENCES

[1] B.Razavi and B.A.Wooley, "Design Techniques for High-Speed High-Resolution Comparators," *IEEE J. Solid-State Circuits*, Vol. SC-27, pp. 1916-1926, Dec 1992.

[2] Y.Wong, M.Cohen, and P. Abshire, "A 1.2-GHz Comparator with adaptable offset in 0.35-um CMOS," *IEEE Trans. Circuits Sys. I,* Vol.55, No.9, Oct. 2008.

[3] D. Fu, K. C. Dyer, S. H. Lewis, and P. J. Hurst, "A digital background calibration technique for time-interleaved analog-to-digital converters," *IEEE J. Solid-State Circuits*, Vol. 33, pp. 1904-1911, Dec. 1998.

[4] J. Doernberg, P. R. Gray, D. A. Hodges, "A 10-Bit 5-Msample/s CMOS Two-Step Flash ADC," *IEEE J. Solid-State Circuits*, Vol. SC-24, pp. 241-249, Apr. 1989.

[5] K. Uyttenhove and M. S. J. Steyaert, "Speed-Power Accuracy Tradeoff in High-Speed CMOS ADCs," *IEEE Trans. Circuits Syst. II*, Vol. 49, pp. 280-287, 2002.

[6] Y. Lin, D. J. Chen and R. Geiger, "Yield enhancement with optimal area allocation for ratio critical analog circuits," *IEEE Trans. Circuits Sys. I*, Vol.53, No.3, March 2006.

[7] L. Sumanen, M.Waltari and K.Halonen, "A mismatch insensitive CMOS dynamic comparator for pipeline A/D converters," in *Proc. IECES*, pp. I-32-35, Dec. 2000.

[8] L. Sumanen, M.Waltari, V.Hakkarainen and K.Halonen, "CMOS dynamic comparators for pipeline A/D converters," in *Proc. ISCAS* pp.V-157-V160, May 2002.

[9] G.A. Al-Rawi, "A New Offset Measurement And Cancellation Technique For Dynamic Latches." *Proceedings of the 2002 IEEE International Symposium on Circuits and Systems*, Vol. V of V, pp. 149-152, May 2002.

[10] J. He, S. Zhan, D. Chen and R. L. Geiger, "A simple and accurate method to predict offset voltage in dynamic comparators," in *Proc. ISCAS*, pp. 1934-1937, May 2008.

[11] P. Nuzzo, F. De Bernardinis, P. Terreni, and G. Van der Plas, " Noise analysis of regenrative comparators for reconfigurable ADC architectures," *IEEE Trans. Circuits Sys. I*, Vol.55, No.6, July 2008.

[12] V. Katyal, R. L. Geiger and D. Chen, "A New High Precision Low Offset Dynamic Comparator for High Resolution High Speed ADCs." *Asia Pacific Conference on Circuits and Systems*, Singapore, pp. 5-8, 2006.

[13] A. Nikoozadeh, B. Murmann, "An Analysis of Latch Comparator Offset Due to Load Capacitor Mismatch," *IEEE Trans. Circuits Syst. II* Express Briefs, Vol. 53, No. 12, December 2006.

[14] T. B. Cho, P. R. Gray, "A 10b, 20 Msample/s, 35mW pipeline A/D converter," *IEEE J. Solid-State Circuits*, Vol. SC-30, no. 5, pp. 166-172, March 1995.

[15] W. Liu, X. Jin, K. M. Cao, and C. Hu, *BSIM4.0.0 MOSFET Model Users' Manual*. Berkeley, CA: Univ. California, 2000.

[16] Y. Chen, C. Hu, *MOSFET Modeling and BSIM3 User's Guide,* Norwell, MA: Kluwer Academic Publishers, 1999.

[17] M.J.M. Pelgrom, A.C.J.Duinmaijer, and A.P.G.Weblbers, "Matching properties of MOS transistors," *EEE J. Solid-State Circuits,* Vol. SC-24, pp. 1433-1439, Oct. 1995.

CHAPTER 3
DETAILED ANALYSES IN PREDICTION OF CAPACITIVE-MISMATCH-INDUCED OFFSET IN DYNAMIC COMPARATORS

A paper accepted by IEEE International Symposium on Circuits and Systems

June 2010

Jun He, Degang Chen, Randall Geiger

ABSTRACT

Due to the positive feedback and the time-varying clock signal, the operating point of each transistor in dynamic comparators is time varying and cannot be analyzed using traditional Op-Amp-based small signal analysis. Until recently, a balanced method is proposed to effectively obtain the analytical models for random offset caused by variations in process parameters. Meanwhile, it has been shown that mismatches from parasitic capacitors are also significant contributors to overall offset. However, the energy storage and nonlinear feature of parasitic capacitor make it even more challenging to analytically predict the capacitive mismatch induced offset. In this work, the previous proposed balance method is generalized and applied to tackle the problem of evaluating capacitive mismatch induced offset. The analytical models are derived to explicitly show offsets caused by capacitor mismatch at different internal nodes. Insights are obtained on identifying the vulnerable nodes to capacitor mismatch and on how to reduce the offset. The numerical example validates the effectiveness of the analytical models.

I. INTRODUCTION

Comparators are used in a wide variety of circuit applications, such as analog-to-digital converter, data transmission, switching power regulator, memories, dynamic logic, sensing amplifier, etc. There are mainly three types of comparators: Op-Amp based comparators, pre-amplifier followed by dynamic latch comparators and pure dynamic comparators [1]. Among the three types, dynamic comparators have the merits of fast speed, zero static power consumption, small area, and therefore serve as fundamental building blocks in high performance analog-to-digital converters (ADCs). Since they are decision-making circuits that compare two analog signals and deliver a logical value at the output, the accuracy, which is often limited by its input referred offset voltage, is essential for the resolution of high performance ADCs.

Neglecting error sources from external circuits, such as timing error of clock signal and variation of reference voltages, the offset voltage in a dynamic comparator is mainly caused by two types of mismatch: (1) static mismatch from variations in process parameters, which are normally dominated by mismatch in μC_{ox} and threshold voltage V_{th}; (2) dynamic mismatch from internal parasitic capacitors' mismatch. In previous work [2], a "balanced method" has been introduced to analytically predict the static offset from variations in process parameters. Most literatures are focusing on reducing the static offset. The methods are proposed, such as increasing *W* and *L* of transistors, and drawing the matching critical pairs as symmetrical as possible [3] [4].

By contrast, dynamic offset caused by mismatch of internal parasitic capacitors is less understood and more challenging to analytically predict. However, offset contributed from internal capacitor mismatch is not trivial. In [5]-[7], the authors pointed out 1fF or

2fF can lead up to several tens of mili-volts offset. In [5], the authors use a simple two-inverter latch as an example to analyze load capacitor mismatch. For other well known dynamic comparator topologies in [4] [8], the same analyses will be too tedious to apply. Also, the model is only for capacitive mismatch at the output node. However, we observed that, besides the output node, some other internal nodes also contribute significant amount of offset. In previous work [9], some preliminary results from simulations are given to show the considerable offset from capacitive mismatch. Design intuitions to reduce offset are obtained from simulations. One model is developed for mismatches at output nodes. In this work, we developed an efficient way to establish the analytical models for capacitive mismatch at different nodes and obtained the insights on how to reduce the capacitive-mismatch-induced offset using the analytical model.

In section II, the proposed "balanced method" for static offset analyses is briefly reviewed. A simple example is given to explain how to apply the same approach to address capacitive mismatch induced offset. In section III, a detailed procedure is given on how to analytically evaluate capacitive mismatch offset in a "Lewis-Gray" structure. In section IV, a numerical example is given to validate the effectiveness of the proposed models. Section V concludes the paper.

II. BALANCED METHOD

Due to internal positive feedback and time-varying clock signal, the operation region of each transistor is time-dependent during each clock period, and how it changes depends on input signal level. Therefore, it is a time-varying non-linear system. In this system, predicting offset becomes very challenging. In [2], we proposed the balance

method to evaluate offset voltage from mismatch in V_{th} and μC_{ox}. The method can also be applied to predict capacitive-mismatch-induced offset.

We would like to introduce the term of balanced mode and the conditions to realize this mode, because they are the core concepts of the proposed balanced method and significantly simplify the analyses of static offset and also capacitive-mismatch-induced offset.

A. Condtions for balanced mode

A typical dynamic comparator [8] as shown in Fig.3.1 can be divided into two halves: left half and right half. Each component in the left half has its counterpart in the right half to form a pair, for instance, M_5 and M_6 form a pair. Under the balanced mode, the two devices within one pair are supposed to be identical in all characteristics including: device type, physics size, process parameters, external bias, operation point, parasitic capacitors, etc. No device mismatch or process variation is presented. Even though dynamic comparators have periodical clock signals and have time-varying operation points, under the balanced mode, the counterparts in each pair will follow the identical time trajectory of operation points as illustrated in Fig.3.1. For instance, time-varying node voltage $V_{out+}(t)$, $V_{s7}(t)$, $V_{s5}(t)$ are the same as $V_{out-}(t)$, $V_{s8}(t)$, $V_{s6}(t)$ respectively through the whole clock cycle, where the subscript *s* means the source of the transistor. Along the time trajectory, we can do linearization and small signal analysis. When the mismatch or small disturbance occurs, ΔV_{in} equal to the input offset voltage is applied to the input terminal to cancel the imbalance effect and to keep the circuit still at or very close to the balanced mode when no mismatch or variation is presented.

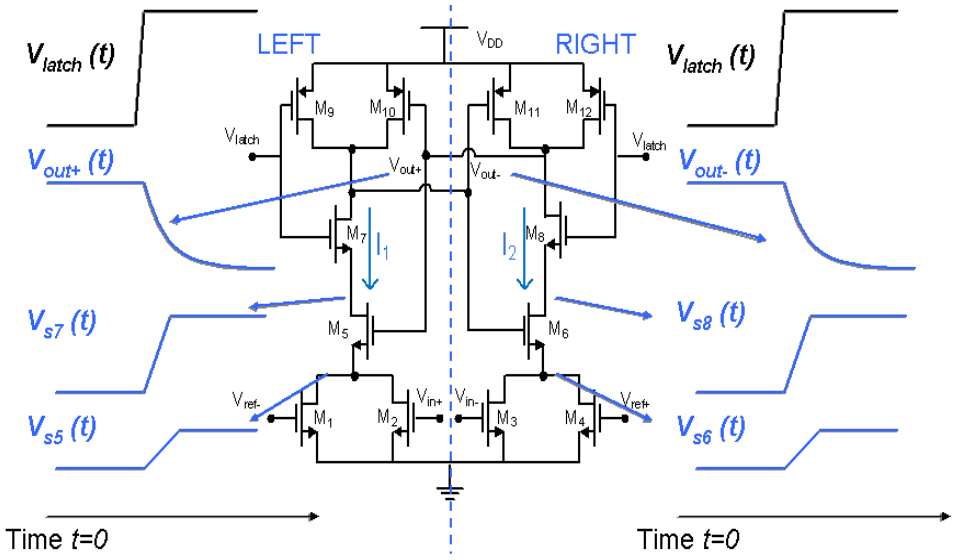

Figure 3.1 A dynamic comparator in balanced mode

B. Capacitive mismatch induced offset prediction using balance mode concept

Capacitive mismatch induced offsets are due to internal parasitic capacitance and are only shown during transient process. A four-terminal MOS device includes twelve different parasitic capacitors [10]. The differential equations with respect to time will get involved to analyze the type of offset because parasitic capacitor is a non-linear and energy storage component. All the above factors make the offset analyses very challenging. The concept of balanced mode can be applied to facilitate the analyses. As the example in Fig.3.2, ΔC demonstrates the overall equivalent mismatch capacitor at the summing drain nodes of M_3 and M_4. Two KCL equations can be expressed as:

$$I_L = I_1 + I_2 + \Delta I + I_{cap\_db1} + I_{cap\_dg1} + I_{cap\_db2} + I_{cap\_dg2} \tag{3.1}$$

$$I_R = I_3 + I_4 + I_{cap\_db3} + I_{cap\_dg3} + I_{cap\_db4} + I_{cap\_dg4} + \Delta C \cdot \frac{dV_{D4}}{dt} \tag{3.2}$$

where I_1, I_2, I_3, I_4 are channel currents through M_1-M_4, respectively. I_{cap_dbi}, I_{cap_dgi}, I_{cap_dbi}, I_{cap_dgi} (i=1, 2, 3, 4) are capacitor charging currents through the parasitic capacitor between two terminals of the transistor M_1-M_4. They can be calculated through the following way:

$$I_{cap\_dbi} = C_{dbi} \cdot \frac{d(V_d - V_b)}{dt} \quad i = 1,2,3,4 \tag{3.3}$$

where d and b mean drain and substrate terminal of a transistor M_i, respectively. I_{cap_dgi}, I_{cap_dbi}, I_{cap_dgi} (i=1, 2, 3, 4) can be found using a similar way.

Under the balanced mode, it is satisfied that:

$$I_L = I_R \,,\, I_1 = I_3 \,,\, I_2 = I_4 \tag{3.4}$$

$$I_{cap\_db1} = I_{cap\_db3} \,,\; I_{cap\_db1} = I_{cap\_db3} \,, I_{cap\_db2} = I_{cap\_db4} \,,\; I_{cap\_dg2} = I_{cap\_dg4} \tag{3.5}$$

Subtract (3.2) from (3.1) and apply the conditions (3.4)-(3.5):

$$\Delta I = \Delta C \cdot \frac{dV_{D4}}{dt} + (I_3 + I_4) - (I_1 + I_2) \approx \Delta C \cdot \frac{dV_{D4}}{dt} \tag{3.6}$$

Since $\Delta I = G_{m2} \cdot \Delta V_{in} + g_{ds2} \cdot \Delta V_{ds} \approx G_{m2} \cdot V_{os}$ (3.7)

The finite output impedance in (3.7) is neglected to simplify the derivation. From (3.6) and (3.7), the offset voltage can be solved as:

$$V_{os} = 1 / G_{m2} \cdot \Delta C \cdot dV_{D4} / dt \tag{3.8}$$

where G_{m2} is the transconductance of M_2. Note that G_{m2} and dV_{D4}/dt have a certain time trajectory within the clock cycle. Here we use G_{m2} and dV_{D4}/dt values under the balanced mode at the moment when V_{latch} is just rising up to V_{DD}. That moment is also supposed to be the time when differential output nodes are released from the reset value V_{DD} and begin to drop. Once the time point is determined, G_{m2}, which is comprised of process

parameters, aspect ratios and bias point, can be calculated using the approach in [2]. The dV_{D4}/dt value can be obtained by measuring the slope of V_{D4} during the short time period right after V_{latch} rises up to V_{DD}.

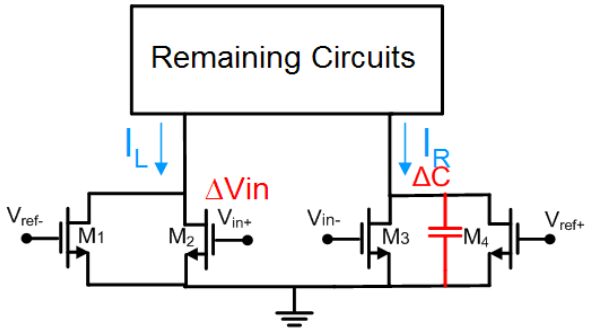

Figure 3.2 Mismatch between internal capacitors

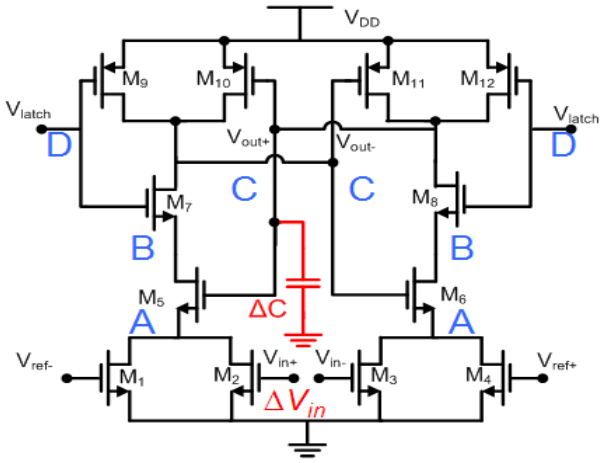

Figure 3.3"Lewis-Gray" structure with an internal parasitic capacitor

III. CAPACITIVE MISMATCH INDUCED OFFSET IN A "LEWIS-GRAY" STRUCTURE

The method in section II can be applied to analyze capacitive mismatch in each internal node in a more complex structure. As "Lewis-Gray" structure shown in Fig.3.3

[8], there has mainly four pairs of internal nodes: *A*, *B*, *C* and *D*. First, consider the capacitive mismatch occurs at differential output nodes V_{out+} and V_{out-} ΔC in Fig.3.4 is the equivalent total mismatch capacitor at the pair of nodes *C*. Assume that there is no other mismatch component. Apply KCL at the source of M_5 and M_6, respectively.

$$\Delta I + I_1 + I_2 + C_{db1} \cdot \frac{dV_{s5}}{dt} + C_{gd1} \cdot \frac{d(V_{s5} - V_{ref-})}{dt} + C_{gd2} \cdot \frac{d(V_{s5} - V_{in+})}{dt}$$

$$= I_5 + C_{bs5} \cdot \frac{d(0 - V_{s5})}{dt} + C_{gs5} \cdot \frac{d(V_{out+} - V_{s5})}{dt} \tag{3.9}$$

$$I_3 + I_4 + C_{db3} \cdot \frac{dV_{s6}}{dt} + C_{gd3} \cdot \frac{d(V_{s6} - V_{in-})}{dt} + C_{gd4} \cdot \frac{d(V_{s6} - V_{ref+})}{dt}$$
$$= I_6 + C_{bs6} \cdot \frac{d(0 - V_{s6})}{dt} + C_{gs6} \cdot \frac{d(V_{out-} - V_{s6})}{dt} \tag{3.10}$$

Under the balanced mode, it is satisfied that:

$$I_1 = I_3 \, , I_2 = I_4 \, , \; I_5 = I_6 \tag{3.11}$$

$$C_{db1} = C_{db3} \, , C_{gd1} = C_{gd3} \, , C_{gd2} = C_{gd4} \, , C_{bs5} = C_{bs6} \, , C_{gs5} = C_{gs6} \tag{3.12}$$

$$V_{ref-} = V_{in-} \, , \; V_{in+} = V_{ref+} \, , \; V_{out+} = V_{out-} \tag{3.13}$$

The condition that $V_{out+} = V_{out-}$ is imposed because when the mismatch component and compensation voltage ΔV_{in} at the input are both present, the circuit is supposed to realize $V_{out+} - V_{out-} = 0$. Subtract (3.10) from (3.9) using the conditions in (3.11)-(3.13):

$$\Delta I \approx -(C_{db1} + C_{gd1} + C_{gd2} + C_{bs5} + C_{gs5}) \cdot \frac{d(V_{s5} - V_{s6})}{dt} \tag{3.14}$$

Then apply KCL at node V_{out+} and V_{out-} respectively.

$$\Delta C \cdot \frac{dV_{out+}}{dt} + C_{gs5} \cdot \frac{d(V_{out+} - V_{s5})}{dt} + I_8 + C_{db8} \cdot \frac{dV_{out+}}{dt} + C_{dg8} \cdot \frac{d(V_{out+} - V_{latch})}{dt}$$
$$= C_{sg10} \frac{d(V_{DD} - V_{out+})}{dt} + C_{db11} \frac{d(V_{DD} - V_{out+})}{dt} + C_{gd11} \frac{d(V_{out-} - V_{out+})}{dt} + C_{gd12} \frac{d(V_{latch} - V_{out+})}{dt} \quad (3.15)$$

$$C_{gs6} \cdot \frac{d(V_{out-} - V_{s6})}{dt} + I_7 + C_{db7} \cdot \frac{dV_{out-}}{dt} + C_{dg7} \cdot \frac{d(V_{out-} - V_{latch})}{dt}$$
$$= C_{sg11} \frac{d(V_{DD} - V_{out-})}{dt} + C_{db10} \frac{d(V_{DD} - V_{out-})}{dt} + C_{gd10} \frac{d(V_{out+} - V_{out-})}{dt} + C_{gd9} \frac{d(V_{latch} - V_{out-})}{dt} \quad (3.16)$$

$$C_{gs5} = C_{gs6}\,,\ C_{db8} = C_{db7}\,,\ C_{dg8} = C_{dg7}\,,\ C_{sg10} = C_{sg11}\,,$$

$$C_{db11} = C_{db10}\ \ C_{gd11} = C_{gd10}\,,\ C_{gd12} = C_{gd9}\,,\ I_8 = I_7\,,\ V_{out+} = V_{out-} \quad (3.17)$$

Subtract (3.16) from (3.15) using the condition (3.17):

$$\Delta C \cdot \frac{dV_{out+}}{dt} = C_{gs5} \cdot \frac{d(V_{s5} - V_{s6})}{dt} \quad (3.18)$$

Combine (3.14) and (3.18):

$$\Delta I \approx \frac{-(C_{db1} + C_{gd1} + C_{gd2} + C_{bs5} + C_{gs5})}{C_{gs5}} \cdot \Delta C \cdot \frac{dV_{out+}}{dt} \quad (3.19)$$

Therefore, the dynamic offset caused by capacitor mismatch ΔC at output node is:

$$|V_{os\_C}| \approx \frac{C_{db1} + C_{gd1} + C_{gd2} + C_{bs5} + C_{gs5}}{C_{gs5} \cdot G_{m2}} \cdot \Delta C \cdot \frac{dV_{out+}}{dt} \quad (3.20)$$

From (3.20), it can be concluded that: first, the analytical model identifies the parasitic capacitors contributing the offset voltage; second, offset is linear proportional to the

mismatch capacitor value ΔC; third, from the term dV_{out+}/dt, it is known that the faster V_{out+} changes since V_{latch} signal is rising up to V_{DD}, the larger offset will be induced.

Figure 3.4 illustrates the most time-efficient way to find dV_{out+}/dt. Under the condition that the circuit is at the balanced mode, V_{out+} or V_{out-} node time trajectory are identical and plotted in Fig.3.4. The estimated dV_{out+}/dt is the slope value measured at the moment when V_{latch} signal is just rising up to V_{DD}. The method only requires one-time short transient analysis (less than one clock cycle) for the circuit at the balanced mode without any mismatch.

Similarly, offset caused by mismatch capacitor ΔC at other pairs of nodes A, B and D can be derived as follows:

$$|V_{os\_A}| \approx \frac{\Delta C}{G_{m2}} \cdot \frac{dV_{s5}}{dt}\ ,\ |V_{os\_B}| \approx \frac{\Delta C}{G_{m2}} \cdot \frac{dV_{s7}}{dt}\ ,\ \ |V_{os\_D}| \approx 0 \quad (3.21)$$

IV. A NUMERICAL EXAMPLE AND SIMULATION RESUTLS

A "Lewis-Gray" structure is implemented in Spectre to validate the effectiveness of the proposed analytical model. A 10fF capacitor is used as the mismatch capacitor and added at nodes *A*, *B*, *C* and *D* respectively. The predicted offsets for each pair of nodes can be calculated using (3.21). Figure 3.5 gives a comparison between predicted values and results from time-consuming transient simulation using a BSIM3v3 model. From Fig. 3.5, they are showing reasonable agreement. The deviation is mainly due to the approximations made, such as the neglect of channel length modulation. The output nodes are the most vulnerable nodes to capacitive mismatch. According to the analytical model (3.20), if the speed requirement can be easily meet, some precisely-matched capacitor can be added simultaneously at V_{out+} and V_{out-} node to reduce dV_{out+}/dt and therefore reduce

offset. Meanwhile, it is still critical to do well-balanced layout including matching of metal routings, via locations and placement of neighboring components so that conditions like (3.17) can be met as close as possible. Nodes *D* are insensitive to capacitive mismatch. Figure 3.6 verifies the observed linear relationship between mismatch capacitor ΔC and offset predicted by (3.20).

Table 3.1 Key values for the dynamic comparator in 0.6µm

Process	**0.6µm CMOS**
Power supply	V_{DD}=3V, V_{SS}=0V
Transistor sizing	$(W/L)_{1,2,3,4}$=16u/1.2u
	$(W/L)_{5,6,7,8}$=16u/1.2u
	$(W/L)_{10,11}$=16u/1.2u
V_{ref}	V_{ref+}=1.6V, V_{ref-}=1.2V
Clock signal V_{latch}	High=3V;Low=0V
	Rise and fall time = 0.3ns
	Pulse width=20ns; Freq=10MHz
Switch (PMOS)	$(W/L)_{9,12}$=16u/1.2u

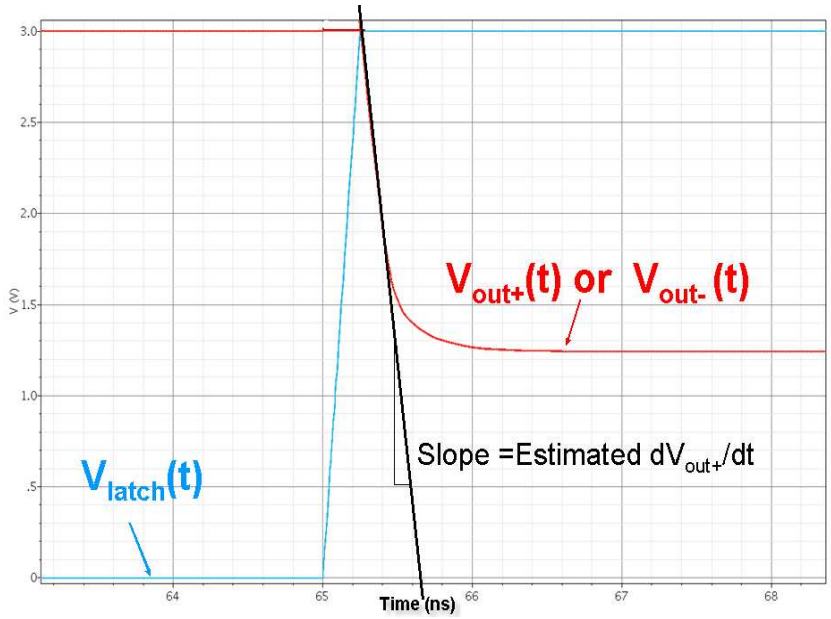

Figure 3.4 Illustration of a method to find the dV_{out+}/dt value

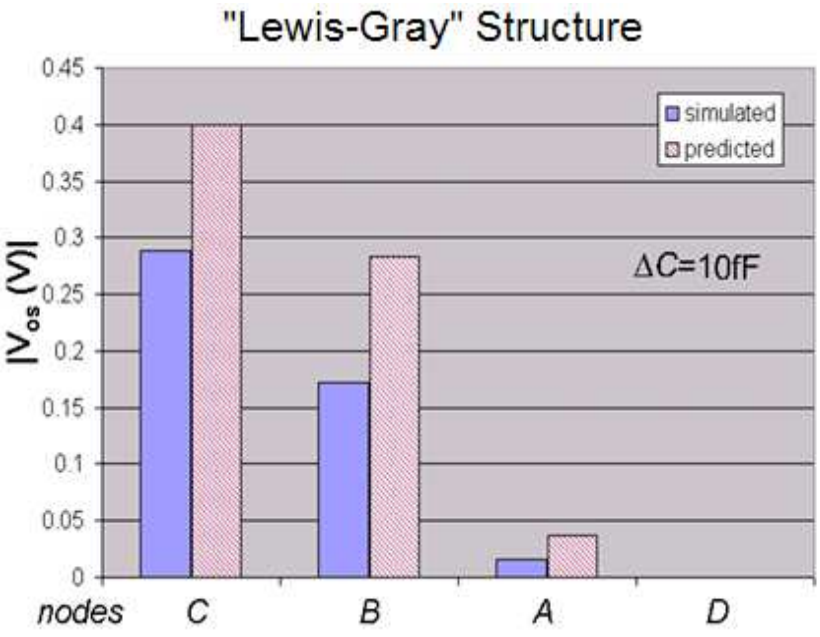

Figure 3.5 Comparison between simulated and predicted offsets

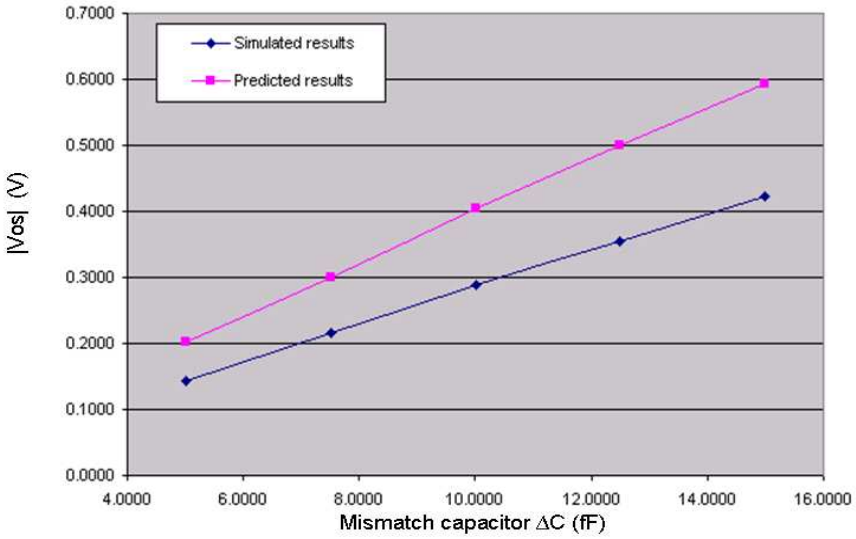

Figure 3.6 V_{OS} vs. capacitor mismatch ΔC at the output nodes

V. CONCLUSIONS

In this work, we apply a "balanced mode" concept to analyze the capacitive mismatch at different nodes in a dynamic comparator. With the aid of the balanced mode concept, the time-varying non-linear problem can be simplified and easily analyzed. The analytical models are derived to demonstrate offsets caused by capacitor mismatch in each pair of nodes. The models correctly identify the nodes that are most vulnerable to capacitive mismatch and offer analytical formulae that provide guides on how to reduce the offset.

REFERENCES

[1] P. E. Allen and D. R. Holberg, *CMOS Analog Circuit Design*, 2nd ed. New York: Oxford Univ. Press, 2002.

[2] J. He, S. Zhan, D. Chen and R. L. Geiger, "A simple and accurate method to predict offset voltage in dynamic comparators," in *Proc. ISCAS*, pp. 1934–1937, May 2008.

[3] B. Razavi and B. A. Wooley, "Design techniques for high-speed high resolution comparators," *IEEE J. Solid-State Circuits*, Vol. 27, no. 6, pp. 1916–1926, Dec. 1992.

[4] L. Sumanen, M. Waltari, and K. Halonen, "A mismatch insensitive CMOS dynamic comparator for pipeline A/D converters," *Proc. IECES*, pp. I-32–35, Dec. 2000.

[5] A. Nikoozadeh and B. Murmann, "An analysis of latch comparator offset due to load capacitor mismatch," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, Vol. 53, no. 12, pp. 1398–1402, Dec. 2006.

[6] V. Katyal, R. L. Geiger, and D. Chen, "A new high precision low offset dynamic comparator for high resolution high speed ADCs," in *Proc. Asia Pacific Conf. Circuits Syst.*, Singapore, pp. 5–8, 2006.

[7] H. M. Attar, M.S. Bakahtiar, "Low voltage low power 8-bit folding/interpolating ADC with rail-to-rail input range," *Journal of Analog Intergrated Circuits and Signal Processing*, Vol.61, No.2, pp.181-189, Nov. 2009.

[8] T. B. Cho and P. R. Gray, "A 10b, 20 Msample/s, 35mWpipeline A/D converter," *IEEE J. Solid-State Circuits*, Vol. 30, no. 3, pp. 166–172, Mar. 1995.

[9] J. He, S. Zhan, D. Chen and R.Geiger, , "Analyses of Static and Dynamic Random Offset Voltages in Dynamic Comparators," *IEEE Transaction on Circuits and* Systems I, Reg. Papers, Vol. 56, No. 5, pp. 911-919, May 2009.

[10] Y. Chen and C. Hu, MOSFET Modeling and BSIM3 User's Guide. Norwell, MA: Kluwer, pp. 37–92, 144–145, 1999.

CHAPTER 4
A DETAIL ANALYSIS OF NONIDEAL EFFECTS ON HIGH PRECISION BANDGAP VOLTAGE REFERENCES

A paper published by 51st Midwest Symposium on Circuits and Systems

pp. 382-385, Aug.2008

Jun He, Randall Geiger, Degang Chen

ABSTRACT

Until recently a closed-form expression for the output voltage of the most basic bandgap references was not available, making it difficult to analytically and systematically determine the effects of the temperature dependence of non-ideal components on the magnitude of the output voltage, on the inflection point location, and on the curvature of these bandgap circuits. In this chapter, several non-ideal components that can adversely affect the performance of bandgap references are identified. A systematical approach is proposed to analytically determine the effects of the temperature dependence of non-ideal components. Analytical expressions for the effects of two of the most common non-ideal components, the temperature-dependent gain-determining resistors and the amplifier offset voltage on the temperature characteristics of basic bandgap circuits are developed. The effectiveness of the analytical expression is proved by comparing with simulation results using Spectre.

I. INTRODUCTION

High precision bandgap voltage references have been required in a wide range of emerging systems, such as high performance data converters, PLLs and monolithic sensors. Despite the performance of reported bandgap voltage references are improving [1,2], little attention has been addressed on theoretical characterization of nonideal effects on bandgap voltage. Not until recently, has an explicit model of the reference involving only process and model parameters been developed to give some insight into how a bandgap operates [3]. However, the explicit model neglects the non-ideal error sources that degrade temperature stability in high precision reference application. A detailed knowledge of error sources' influences on bandgap behavior, such as on the inflection point, the curvature of the bandgap curve and the value of the reference output, is fundamental in affording the designers a better understanding of main limitations and to improve the approach to design high precision references.

In this chapter, we apply Kujik's bandgap voltage reference topology [4] because it represents the basis for other new types of structures. The non-ideal error sources are illustrated in Fig.4.1 including temperature-dependent offset voltage $V_{os}(T)$, temperature dependence of resistors $R_0(T)$~$R_2(T)$, matching between $R_0(T)$~$R_2(T)$, matching between diodes D_1 and D_2, finite gain of op amp A(s), parasitic resistors in pn junctions $R_{d1}(T)$, $R_{d2}(T)$. Besides, V_{GO} temperature dependence [5], error from package stress [6] will also introduce more errors. This work will focus on the effects of temperature dependence of gain-determining resistors $R_0(T)$~$R_2(T)$ and offset voltage of op amp $V_{os}(T)$. Almost all the authors simply gloss over the fact of the temperature dependence of R_0~R_2, because in the ideal situation R_0~R_2 always appear in a form of resistor ratio, and temperature

coefficients (TCs) of R_0~R_2 can be canceled out and will not affect the thermal stability of reference output voltage. However, the explicit closed form model in [3] indicates that a single resistor value not just resistor ratios will also appear in the expression of $V_{ref}(T)$. Offset voltage $V_{os}(T)$ is the biggest error source that causes the non-reproducibility in the output voltage temperature coefficient [1].

In section II, a systematic approach to derive an explicit model will be introduced. The effects of temperature-dependent gain-determining resistors and offset voltage of op amps are both included into the explicit model. Their influences on the inflection point, the curvature of the bandgap curve and the value of the bandgap output are discussed separately. In section III, Spectre simulation results are compared with theoretical analysis from the model and show a good consistency of the analysis. Section IV concludes the work.

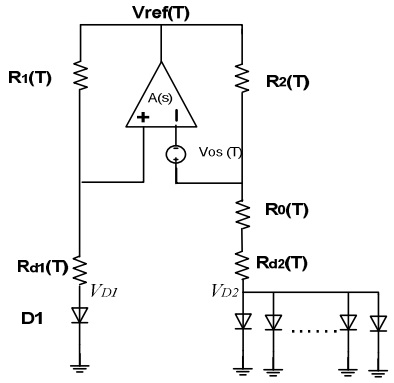

Figure 4.1 A bandgap circuit with two error sources

II. DETAILED CHARACTERIZATIONS OF BANDGAP REFERENCES WITH NON-IDEALITIES

In the circuit in Fig.4.2, assume that R_0~R_2 have a certain temperature coefficient TCR, so they are modeled by

$$R_i(T) = R_{iN} \cdot (1 + TCR \cdot (T - T_0))_{i=0,1,2} \tag{4.1}$$

where R_{iN} is the nominal value of R_0~R_2 at temperature T_0 . Also assume that input referred offset voltage V_{os} has the first order temperature coefficient TCV_{os} and can be modeled as

$$V_{os}(T) = V_{os}(T_0) + TCV_{os}(T - T_0) \tag{4.2}$$

Five equations can be determined to completely characterize the circuit in Fig.4.2.

$$V_{ref} = I_{D1}R_1(T) + V_{D1} \tag{4.3}$$

$$V_{D1} = V_{D2} + I_{D2}R_0(T) - V_{os}(T) \tag{4.4}$$

$$I_{D2} = [V_{ref} - V_{D1} - V_{os}(T)]/R_2(T) \tag{4.5}$$

$$V_{D1} = V_t \ln(I_{D1}) + V_{GO} - V_t[\ln(J_{sx}A_1) + m \ln T] \tag{4.6}$$

$$V_{D2} = V_t \ln(I_{D2}) + V_{GO} - V_t[\ln(J_{sx}A_2) + m \ln T] \tag{4.7}$$

where J_{sx} are process parameters and independent of temperature. m is also process parameter equal to 2.3. V_{GO} is bandgap voltage 1.205V. A_1 and A_2 are the area factors for diodes D_1 and D_2 respectively. $Vt=kT/q$, k is Boltzman's constant, T is temperature in K, and q is the charge of an electron.

The set of equations (4.3)-(4.7) in the unknowns $\{I_{D1}, I_{D2}, V_{D1}, V_{D2}, V_{ref}\}$ completely characterize the circuit in Fig.4.2. After eliminating V_{D1}, V_{D2}, and I_{D2}, we reduce (4.3)-(4.7) to a set of 2 equations in the unknowns $\{I_{D1}, V_{ref}\}$.

$$V_{ref} = I_{D1}R_1(T) + V_t \ln(I_{D1}) + V_{GO} - V_t[\ln(J_{sx}A_1) + m \ln T] \tag{4.8}$$

$$I_{D1}R_1(T) = \frac{R_2(T)}{R_1(T)} V_t \ln\left(\frac{I_{D1}R_2(T)}{I_{D1}R_1(T) - V_{os}(T)}\right) + \frac{R_2(T)}{R_0(T)} V_t \ln\left(\frac{A_2}{A_1}\right) + V_{os}(T) \cdot \left(1 + \frac{R_2(T)}{R_0(T)}\right) \tag{4.9}$$

These two equations are highly non-linear and highly coupled. We will first linearize (4.9) from which I_{D1} can be obtained. The first logarithmic term in (4.9) can re-arranged as

$$\ln\left(\frac{I_{D1}R_2(T)}{I_{D1}R_1(T) - V_{os}(T)}\right) = \ln\left[\frac{R_2(T)}{R_1(T)} \cdot \left(\frac{1}{1 - \frac{V_{os}(T)}{I_{D1}R_1(T)}}\right)\right] \tag{4.10}$$

From (4.10), note that $I_{D1}R_1(T)$ is the voltage drop over R_1, and also the voltage difference between bandgap reference output voltage V_{ref} and diode voltage V_{D1}. Normally, V_{ref} is around 1.2V and V_{D1} is around 0.6V, so

$$I_{D1}R_1(T) = V_{ref} - V_{D1} >> V_{os}(T) \tag{4.11}$$

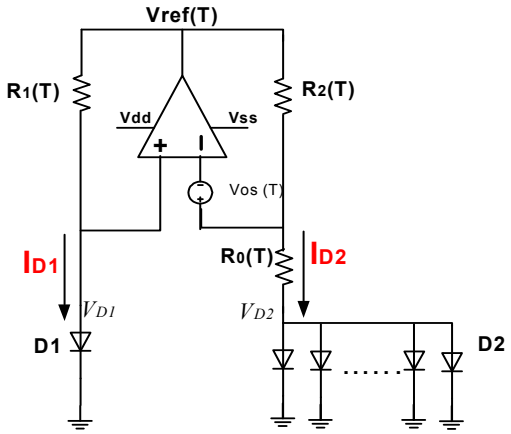

Figure 4.2 A bandgap circuit with temperature-dependent R_0~R_2 and Vos(T)

Applying the condition in (4.11), (4.10) can be expressed as

$$\ln[\frac{R_2(T)}{R_1(T)}\cdot(\frac{1}{1-\frac{V_{os}(T)}{I_{D1}R_1(T)}})]=\ln\frac{R_2(T)}{R_1(T)}+\ln(\frac{1}{1-\frac{V_{os}(T)}{I_{D1}R_1(T)}})$$

$$\approx \ln\frac{R_2(T)}{R_1(T)}+\frac{V_{os}}{I_{D1}R_1(T)}\approx \ln\frac{R_2(T)}{R_1(T)} \tag{4.12}$$

With the approximations, we have linearized (4.10) as

$$I_{D1}R_1(T)=\frac{R_2(T)}{R_1(T)}V_t\ln(\frac{R_2(T)}{R_1(T)})+\frac{R_2(T)}{R_0(T)}V_t\ln(\frac{A_2}{A_1})+V_{os}(1+\frac{R_2(T)}{R_0(T)}) \tag{4.13}$$

Eliminating I_{D1}, finally gives a general close form explicit expression for V_{ref} as

$$V_{ref}=a+bT+cT\ln(T)+d\ V_{os}(T) \tag{4.14}$$

where $a=V_{G0}$ (4.15)

$$b=\frac{k}{q}(\frac{R_2(T)}{R_0(T)}\ln(\frac{R_2(T)}{R_1(T)}\frac{A_2}{A_1})+\ln(\frac{R_2(T)}{R_1(T)}\frac{1}{R_0(T)}\frac{\ln(\frac{R_2(T)}{R_1(T)}\frac{A_2}{A_1})}{A_1J_{sx}})) \tag{4.16}$$

$$c=\frac{k}{q}(1-m) \tag{4.17}$$

$$d=1+\frac{R_2(T)}{R_0(T)}+\frac{(1+\frac{R_2(T)}{R_0(T)})}{R_2(T)/R_0(T)\cdot\ln(R_2(T)/R_1(T)\cdot A_2/A_1)} \tag{4.18}$$

Non-ideal effects from Temperature dependent offset voltage V_{os} (T) and gain-determining resistor R_0(T)~R_2(T) are inherently involved in (4.14). Inflection point, the curvature of the bandgap curve and the value of the bandgap output are readily to be analyzed.

Temperature dependence of R_0~R_2

First, investigate non-ideal effect from temperature dependent gain-determining resistor $R_0(T)$~$R_2(T)$ by assuming that $V_{os}(T)=0$. Equation (4.14) can be simplified as

$$V_{ref} = a + bT + cT\ln(T) \tag{4.19}$$

Note that resistors are not always shown in a form of a resistor ratio. In (4.16), $R_0(T)$ appear as a single resistor value dependent on temperature variation, which intuitively explains why temperature dependence of $R_0(T)$~$R_2(T)$ is necessary to be considered.

According to (4.1), *b* in (4.16) can be re-written as

$$b = b_1 = \frac{k}{q}\left(\frac{R_{2N}}{R_{0N}}\ln\left(\frac{R_{2N}}{R_{1N}}\frac{A_2}{A_1}\right) + \ln\left(\frac{R_{2N}}{R_{1N}}\frac{1}{R_0(T)}\frac{\ln\left(\frac{R_{2N}}{R_{1N}}\frac{A_2}{A_1}\right)}{A_1 J_{sx}}\right)\right) \tag{4.20}$$

Based on (4.19), (4.15), (4.17) and (4.20), inflection point analysis can be done by differentiating V_{ref} with respect to *T* and setting the derivative equal to 0 gives

$$\frac{\partial V_{ref}}{\partial T} = b_1 + T\frac{\partial b_1}{\partial T} + c\ln(T) + c = 0 \tag{4.21}$$

From (4.20), differentiate b_1 with respect to T gives

$$\frac{\partial b_1}{\partial T} = \frac{k}{q}R_{0N}\cdot TCR\cdot\frac{-1}{R_0(T)} \approx -\frac{k}{q}TCR \tag{4.22}$$

Substitute (4.22) into (4.21) to solve for T, and T is the new inflection point T_{infN1} when TC of R_0~R_2 is considered. Equation (4.21) is re-arranged as

$$b_1 - T_{\inf N1}\cdot\frac{k}{q}\cdot TCR + c\ln(T_{\inf N1}) + c = 0 \tag{4.23}$$

It has been derived in [3] that if no error sources are considered, the desired inflection point $T_{inf I}$ satisfies the following equation

$$b_I + c_I + c_I \ln T_{\inf I} = 0 \tag{4.24}$$

where $b_I \approx b_1, c_I = c$.

Take the difference between (4.23) and (4.24) and rearrange the expression. It gives

$$T_{\inf I} = T_{\inf N1} \cdot \exp(T_{\inf N1} \cdot \frac{TCR}{m-1}) \tag{4.25}$$

From (4.25), it can be concluded that the new inflection point T_{infN1} is readily to be solved when *TCR* is known.

By doing the second derivative of the reference voltage evaluated at the inflection point, it can be determined that how rapidly the reference curve opens up away from the inflection point. It is also called the curvature of the reference curve. According to (4.21) and (4.22), it follows that

$$\frac{\partial^2 V_{ref}}{\partial T^2}|_{T=T\inf N1} = \frac{\partial b_1}{\partial T} + (\frac{\partial b_1}{\partial T})^2 + \frac{c}{T_{\inf N1}} = (-\frac{k}{q}TCR) + (-\frac{k}{q}TCR)^2 + \frac{c}{T_{\inf N1}} \tag{4.26}$$

Offset voltage of op amp and its temperature dependence

When the effect from offset voltage $V_{os}(T)$ is considered, assume all the other non-ideal effects are not existing. *b* and *d* in the general expression can be rewritten as

$$b = b_2 = \frac{k}{q}(\frac{R_{2N}}{R_{0N}}\ln(\frac{R_{2N}}{R_{1N}}\frac{A_2}{A_1}) + \ln(\frac{R_{2N}}{R_{1N}}\frac{1}{R_{0N}}\frac{\ln(\frac{R_{2N}}{R_{1N}}\frac{A_2}{A_1})}{A_1 J_{sx}})) \tag{4.27}$$

$$d = d_2 = 1 + \frac{R_{2N}}{R_{0N}} + \frac{(1 + R_{2N}/R_{0N})}{R_{2N}/R_{0N} \cdot \ln(R_{2N}/R_{1N} \cdot A_2/A_1)} \tag{4.28}$$

It can be seen that b_2 and d_2 are independent of temperature. Then apply the same method to (4.14) as in part *A* to analyze the inflection point. It follows that

$$\frac{\partial V_{ref}}{\partial T} = b_2 + c\ln(T) + c + d_2 \frac{\partial V_{os}(T)}{\partial T} = 0 \tag{4.29}$$

This expression can be solved for T to determine the inflection point T_{infN2} to be

$$T_{\inf N2} = \exp[-b_2/c - 1 - (d_2/c_2)\frac{\partial V_{os}(T)}{\partial T}]$$

$$= \exp(-b_2/c - 1) \cdot \exp[-(d_2/c_2)\frac{\partial V_{os}(T)}{\partial T}] = T_{\inf I} \cdot \exp[-(d_2/c_2) \cdot TCV_{os}] \tag{4.30}$$

The curvature around the inflection point is calculated as

$$\frac{\partial^2 V_{ref}}{\partial T^2}|_{T=T\inf N2} = \frac{c}{T_{\inf N2}} \tag{4.31}$$

Comparison with bandgap reference characteristics with no error sources

As it is seen from the above obtained inflection points, curvature and V_{ref} magnitude information, it is worthy making a comparison between ideal characteristics and the ones with error sources V_{os} (T) and temperature-dependent R_0~R_2. The results are listed in Table 4.1.

Table 4.1 Comparison between different conditions

Conditions	Characteristics Comparison		
	Inflection Point (T_{inf})	Curvature around T_{inf}	$V_{ref}(T_{inf})$
Ideal Transfer Curve	$T_{inf1}=\exp(-b/c-1)$	c/T_{inf1}	$a+b\cdot T_{inf1}+c\cdot T_{inf1}\cdot\ln(T_{inf1})$
Transfer curve with Vos	$T_{inf1}\exp(-d/c\cdot TCV_{os})$	$c/[T_{inf1}\exp(-d/c\cdot TCV_{os})]$	$a+b\cdot T_{infN2}+$ $c\cdot T_{infN2}\cdot\ln(T_{infN2})+d\cdot Vos(T_{infN2})$
Transfer curve with $R_0(T)-R_2(T)$	$f(T_{inf1}, TCR, m)=T_{infN1}$	$(-k/q\cdot TCR)+(k/q\cdot TCR)^2$ $+c/T_{infN1}$	$a+b\cdot T_{infN1}+c\cdot T_{infN1}\cdot\ln(T_{infN1})$

III. NUMERICAL EXAMPLES AND SPECTRE SIMULATION

Kujik's circuit in Fig.4.2 is implemented in AMI0.6μ process in Cadence to verify the effectiveness of derivation in previous section. The key values are as follows: V_{dd}=5V, V_{ss}=0V, $R_{2N}=R_{1N}$=5.95KΩ, R_{0N}=786Ω, A2:A1=8:1, A(s=0) =100dB, $V_{os}(T_0=27^oC)$=1mV. The effect of temperature dependent resistor $R_0(T)\sim R_2(T)$ on the inflection point T_{inf}, the curvature near T_{inf} and the value of the bandgap output V_{ref} at T_{inf} is simulated, and then compared with the calculated values that can be found from explicit expression in (4.19), (4.25) and (4.26). It is worth mentioning that in the simulation, T_{inf} and V_{ref} at T_{inf} can be easily determined by measuring V_{ref} peak location in the plot of V_{ref} vs. temperature transfer curve. To get the curvature near T_{inf}, quadratic curve fitting can be first applied to fit the transfer curve near T_{inf} by a parabola, and the curvature is the second order coefficient. From Fig.4.3-4.5, the calculated values from the explicit model are in close

agreement with the simulated values. The differences are mainly due to the approximation made during the derivation, the model parameters and the I_C-V_D diode characteristics.

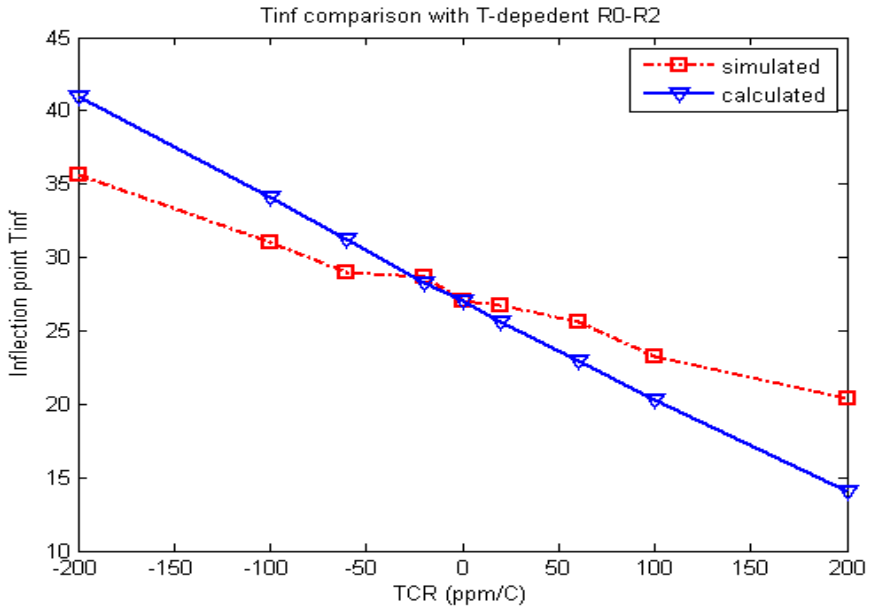

Figure 4.3 Inflection point vs. temperature coefficient of R_0~R_2 (TCR)

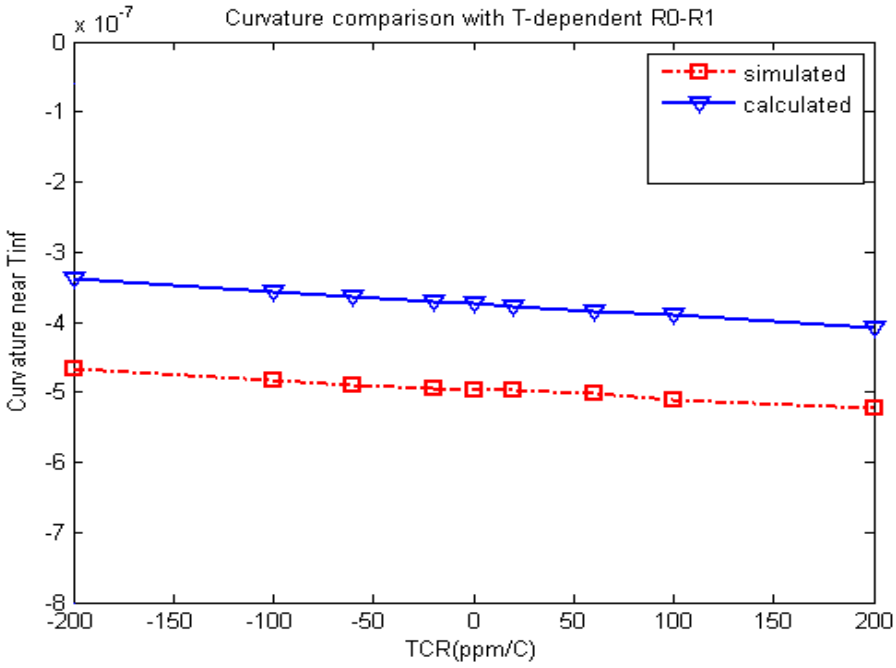

Figure 4.4 Inflection curvature near T_{inf} vs. TCR

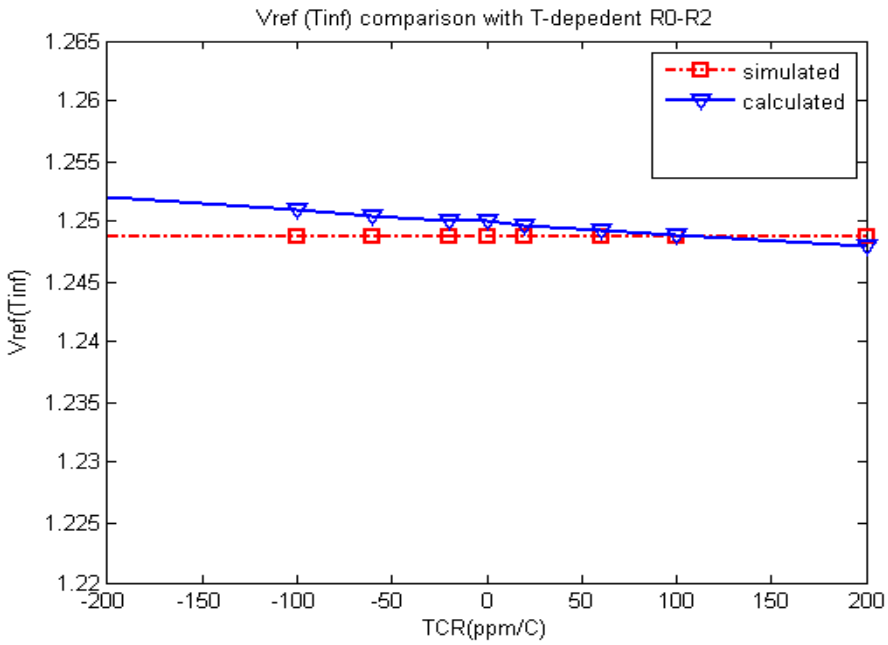

Figure 4.5 V_{ref} value near T_{inf} vs. TCR

IV. CONCLUSIONS

A systematical approach is proposed in this work to analytically determine the effects of the temperature dependence of non-ideal components on the inflection point location, on the curvature of bandgap curve and on the magnitude of the output voltage. The effects of two of the most common non-ideal components, the temperature-dependent gain-determining resistors and the amplifier offset voltage V_{os} (T), on the temperature characteristics of basic bandgap circuits are analyzed. The effectiveness of the derived model is shown by comparing with simulation results using BSIM3v3 model in Spectre. This new approach can allow the circuit designers to have a better understanding of main limitations of the adopted voltage references and to improve the approach to design high precision reference circuits.

REFERENCES

[1] B.-S. Song and P. R. Gray, “A precision curvature-compensated CMOS bandgap reference,” *IEEE J. Solid-State Circuits*, Vol. SC-18, pp. 634–643, Dec. 1983.

[2] I. Lee, G. Kim, and W. Kim, “Exponential curvature-compensated BiCMOS bandgap references,” *IEEE J. Solid-State Circuits*, Vol. 29, pp. 1396–1403, Nov. 1994.

[3] X. Dai, D.Chen and R. Geiger, “Explicit characterization of bandgap references,” in *Proc. ISCAS*, 4 pp.573-576, May 2006.

[4] K. Kuijk, “A Precision Reference Voltage Source”, *IEEE Journal of Solid State Circuits*, Vol. 8, pp. 222-226, June 1973.

[5] Y. P. Tsividis, "Accurate analysis of temperature effects in IC VBE characteristics with application to bandgap reference sources," IEEE J. Solid-State Circuit, Vol. SC-15, pp. 1076-1084, Dec. 1980.

[6] G.A. Rincon-Mora, *Voltage References*, IEEE Press, John Wiley & Sons, Inc., 2002.

Appendix to Chapter 4: Laboratory Measurement Results and Discussions on Positive Curvature Bandgap Voltage Reference

I. INTRODUCTION

The bandgap voltage of silicon is a physical property that is essentially independent of temperature. Good bandgap circuits present, at an output, a voltage that is dominantly dependent on the bandgap voltage, and thus one that is very insensitive to temperature variations, process parameters, and power supply variations.

Bandgap references have been widely used since the mid 1970's. The first bandgap reference was introduced by Hilbiber in 1964 [1]. A lot of pioneering work on bandgap references is attributable to Widlar [2]-[3], who published some of the first papers on the subject in 1969 and 1971 while working at National Semiconductor. Another early paper on the topic is that of Brokow [4] in 1974, who was working at Analog Devices at the time. The basic bandgap references derived their name from the observation that the output voltage of such references is dependent upon the bandgap voltage of silicon. Most basic bandgap references reported have a voltage-temperature relationship that has a single inflection point, although some of voltage references may have two or more inflection points or may use other forms of curvature compensation. Those structures that have two or more inflection points are invariably more complicated but presumably the temperature effects on the output voltage will be less significant. From the previous comments, it is seen that thermal stability is one of the most important characteristics of a voltage reference. The temperature coefficient (TC) is often used to characterize the thermal stability of a reference and is defined by:

$$TC = \frac{10^6}{2} \cdot \frac{V_{\max} - V_{\min}}{V_{nom}(T_{\max} - T_{\min})} \text{ppm} / {}^{o}C \tag{4.32}$$

where V_{min}, and V_{max} are the minimum and maximum voltage levels over the temperature range from T_{min} to T_{max}, respectively and V_{nom} is the nominal voltage value at which the circuit is intended to be operated.

Due to the widespread use of voltage references as the reference for data converters, it is often convenient to express the TC in terms of the resolution needed for data converters. If the data converter has a resolution of *n* bits and a total error of 1 LSB over temperature in the reference can be tolerated, it follows from (4.32) that the TC can be expressed in terms of the resolution of the data converter, it can support by using the equation:

$$TC = \frac{10^6}{2^{n+1}} \cdot \frac{1}{T_{\max} - T_{\min}} \text{ppm} / {}^{o}C \tag{4.33}$$

Table 4.2 provides the mapping between data converter resolution and the required voltage reference TC.

Table 4.2 TC requirements for data converters

bits	TC for 0~70°C Commercial range	TC for 0~100°C Commercial Range	TC for -55~125°C Military
8	27.9	19.5	10.9
9	14.0	9.8	5.4
10	7.0	4.9	2.7
11	3.5	2.4	1.36
12	1.7	1.22	0.68
13	0.87	0.61	0.34
14	0.44	0.31	0.17
15	0.22	0.15	0.085
16	0.11	0.076	0.042

Since the work of Widlar in the seventies, considerable effort has been made on improving the performance of basic bandgap circuits to meet the ever-increasing temperature stability requirements or specific needs from high precision applications. One of the most recent works is that of Rincon-Mora and Allen [5]. They used a higher-order curvature compensation network and required resistor trimming throughout the temperature range of the circuit. With external trimming by hand, they reported a TC of ±10ppm/°C over a 105°C temperature range for a 0.6 V reference in a basic CMOS process. A comparison with Table 4.2 shows this is at approximately the 9-bit level. Buck et al. [6] used a more standard single inflection point bandgap structure in a standard CMOS process but instead of trimming with resistors, they hand-trimmed current sources (presumably at multiple temperatures, although the trimming procedure was not discussed) by digitally switching in and out small transistors. Their nominal reference voltage was a standard 1.2 V. They reported a best case TC over 25 samples of ±64 ppm/°C over a limited 70°C operating range. Amena [7] from Phillips in Eindhoven focused on low voltage operation in a standard CMOS process. He used a standard single inflection point bandgap reference. He reported on a 0.65 V trimmed structure with a TC of ±29 ppm/ºC operating over a 120ºC temperature range but did not discuss any details about trimming. In the same paper, he reported a high-precision, non-trimmed 1.2V structure with a TC of ±6.7ppm/°C over the same 120°C operating range. A comparison of these results with Table 4.2 shows that this performance is at approximately the 9.5 bit level. This performance is more in line with some commercial products but it is worth noting that Amena has 10 years of experience in an industrial design environment at Phillips. Banba et al. from Toshiba in Japan [8] discussed their results with very sketchy

details about the circuit and the experimental results. They indicated that test equipment measurement accuracy was limited to 1mV, and with their 500mV reference voltage over a 100ºC operating range, they would have a measurement uncertainty of ±20 ppm/ºC. Their circuit does include matching critical resistors but the issue of trimming was not discussed. Based upon their pin count, it is presumed that their circuit was not trimmed. From data reported in the paper, an inferred TC of ±60 ppm/ºC over a 98°C operating range appears to have been obtained but with their reported uncertainty in measurements, the actual performance is difficult to ascertain.

In this appendix, a model for the temperature characteristics of a base-emitter formed pn junction will be introduced, since it forms the basis of a bandgap voltage reference. Using this model for the pn junction, it will be shown that the basic bandgap voltage reference structures have a negative curvature near the inflection point. This observation has been seen from simulations and is often reported with measurement results that have appeared in the literatures. However, our measurement results for an implementation of Kujik's bandgap reference circuit show a measured curvature with an opposite sign from that reported for the single inflection point bandgap references [2]–[4], [6]–[8], [10]. The measured temperature coefficient is at the 7~15 ppm/^{o}C level from three chip samples with the integrated diodes coming from a standard 0.6μm process.

II. CURVATURE OF BANDGAP VOLTAGE REFERENCE WITH SINGLE INFLECTION POINT

The relationship between diode current and voltage drop for a pn junction is given by the well-known exponential relationship:

$$I_C = I_S \cdot \exp(\frac{V_D}{n \cdot V_t}) \tag{4.34}$$

where I_s is saturation current and n is a slope factor that is approximately equal to 1 [12]. V_T is the thermal voltage equal to $k \cdot T/q$, where k is Boltzmann's constant and q is the charge of an electron. The saturation current is also a temperature dependent quantity, which can be expressed as:

$$I_C = J_{sx} \cdot A \cdot (T^m \cdot \exp(-\frac{V_{GO}}{V_t})) \cdot \exp(\frac{V_D}{n \cdot V_t}) \tag{4.35}$$

where A is the diode area factor, J_{sx} is a process parameter, and m is process parameter approximately equal to 2.3.

The pn junction forms the basic building block of voltage references. The voltage drop over one diode, like V_{D1} in Fig.5.1, has a negative temperature coefficient. If the relative currents are appropriately selected, the difference between two diodes voltages V_{D1}-V_{D2}, has a positive temperature coefficient. Bandgap references are designed so that the output voltage V_{ref} is the weighted addition of these two signals, where the weight is chosen such that the sum has a zero temperature coefficient at the desired temperature.

Neglecting error sources, such as the operational amplifier offset voltage, the temperature coefficient of the resistors, the parasitic resistance in the interconnect, and well base spreading resistance, the output voltage for the circuit in Fig.4.6 [10] can be described by the following expressions using the systematic approach discussed in Chapter 4:

$$V_{ref} = a + bT + cT\ln(T) \tag{4.36}$$

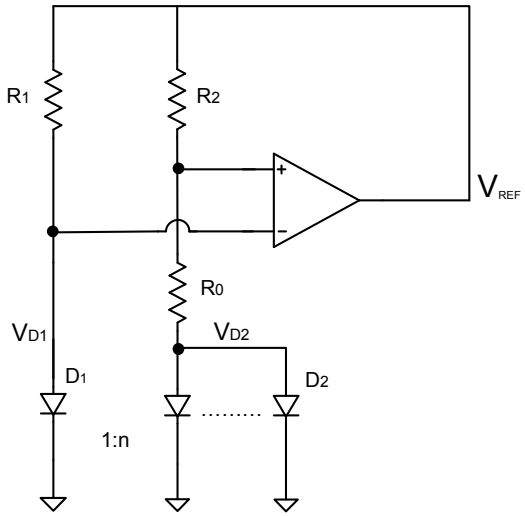

Figure 4.6 Circuit diagram for laboratory measurement

where $a = V_{G0}$ (4.37)

$$b = \frac{k}{q}\left(\frac{R_2}{R_0}\ln\left(\frac{R_2}{R_1}\frac{A_2}{A_1}\right) + \ln\left(\frac{R_2}{R_1}\frac{1}{R_0}\frac{\ln\left(\frac{R_2}{R_1}\frac{A_2}{A_1}\right)}{A_1 J_{sx}}\right)\right) \tag{4.38}$$

$$c = \frac{k}{q}(1-m) \tag{4.39}$$

where A_1 and A_2 are the area factors for diodes D_1 and D_2, respectively.

The second derivative of the reference voltage evaluated at the inflection point is a measure of how rapidly the reference curve opens up away from the inflection point. It is also called the curvature of the reference curve. Upon taking the second derivative, it follows that the curvature for the basic bandgap reference can be expressed as:

$$\frac{\partial^2 V_{ref}}{\partial T^2}\Big|_{T=T\,\mathrm{inf}} = \frac{c}{T_{\mathrm{inf}}} = (1-m)\frac{k}{q}\frac{1}{T_{\mathrm{inf}}} \tag{4.41}$$

Since *m* is a parameter larger than 1, the curvature of this voltage reference is expected to have a negative sign. This characteristic has been observed for years in the literatures. The typical bandgap voltage output should have a transfer curve as illustrated in Fig. 4.7.

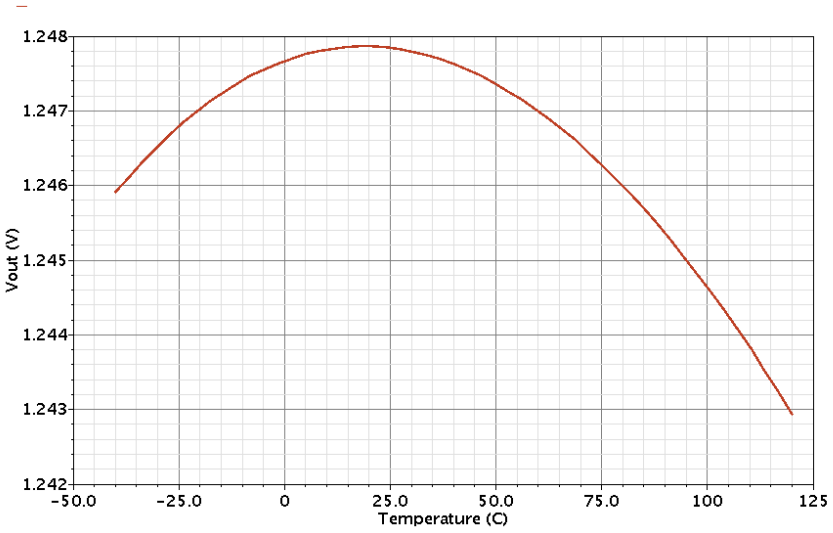

Figure 4.7 Typical bandgap voltage reference transfer curve

III. LABORATORY MEASUREMENT RESULTS AND DISCUSSIONS

The circuit in Fig.4.6 was built using on-chip diodes, an external ua741 general purpose operational amplifier [14], external thin film resistors, a solderless breadboard and jumper wires. The on-chip diodes were implemented in 0.6µm process, and the layout is shown in Fig.4.8. The two diodes D_1 and D_2 have an area ratio of 1:8. The green layer is nwell forming the cathode of the diode with the dimension 70.8µm x 83.4µm. The anode is formed by a array of 3.9µm x 3.9µm Metal 1 to P+ diffusion via. The electrical connection for the anode to pad is from Metal 1 to Metal 2 via and then to Metal 2, and the electrical connection for the cathode is from nwell to Metal 1. The detailed dimension information is shown in Fig.4.8, and in Fig.4.9 showing a magnified

layout for diode D_1. The gain-determining resistors R_1 and R_2 are designed to be the same, but due to the process variation, R_1 is 5.922 KΩ and R_2 is 5.919 KΩ when implementing the circuit using the thin film resistors in the lab. R_0 is 779 ohms.

In the initial measurements, the whole circuit was placed in an environmental oven (chamber). A Tenny Jr. environmental chamber with manual temperature control was used for these measurements. The overall measurement time was rather long since it took considerable time (approximately 2 hours) for the temperature of the test circuits inside the chamber to reach the thermal equilibrium if a significant change in temperature was made. Since the chamber did not have computer control capabilities, personal intervention was required for every change in temperature. The measured output voltage transfer curve is shown in Fig.4.10. The temperature range is approximately -10^{o}C to 60^{o}C with 10^{o}C measurement steps. At each temperature point after thermal equilibrium was reached, a Labview controlled multi-meter took 100 sample points at 1s intervals, denoted as samples of V_{ref}. The error bar plot in Fig.4.10 demonstrates the mean value and the 3σ value of the 100 samples at each temperature point. The measured transfer curve of V_{ref} shows a different characteristic in curvature from that of the usual bandgap voltage references in Fig.4.7. The temperature coefficient is 7.3ppm/^{o}C over a 65^{o}C temperature range. A comparison with Table 4.2 shows that the performance is at approximately the 10-bit level.

From a theoretical analysis, neglecting offset voltage effects, temperature dependence of the resistors, and assuming the gain of the op-amp and the β of the transistors is very large, it can be shown that the TC will be negative and is around -2.5 ppm/^{o}C. However, reported measurements of TC of circuits such as this are typically

around 15.7 ppm/^{o}C [3]. The measured TC of 7.3ppm/^{o}C would be considered a very good performance.

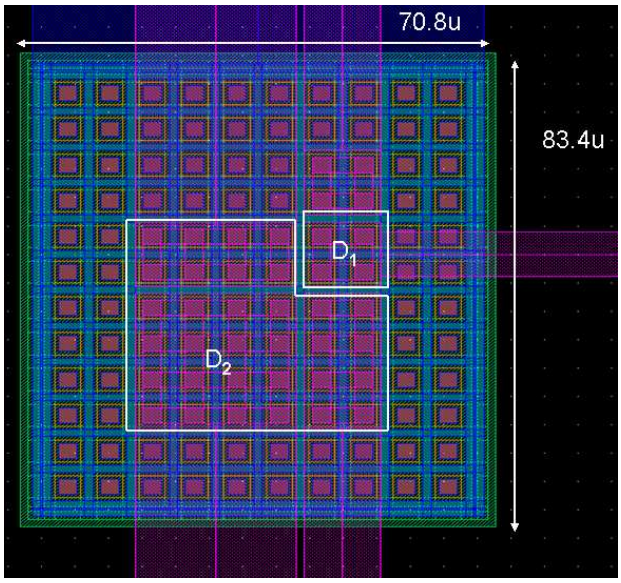

Figure 4.8 Layout of on-chip diodes

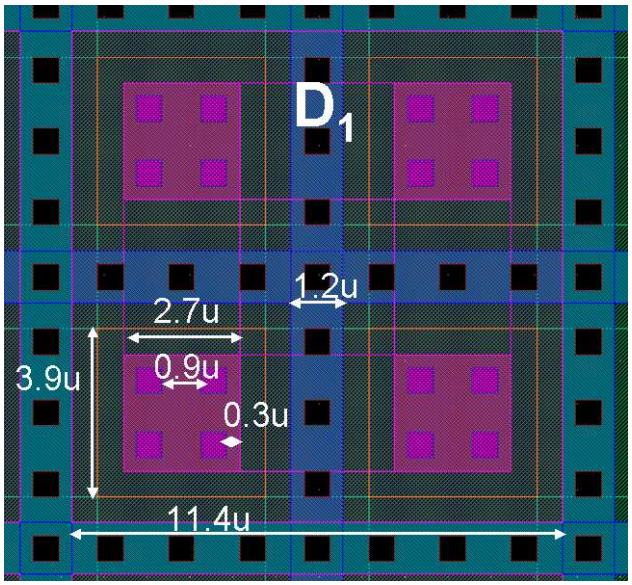

Figure 4.9 Zoomed-in layout of on-chip diode D_1

Since the curvature is quite different from all other reported single inflection point bandgap voltage references [2]–[4], [6]–[8], [10], we have tried to identify the reason. First, the ua741 op-amp was replaced by an ultra low offset op amp, LTC1052 [13] to

eliminate the possible influence from offset voltage. The measured transfer function is shown in Fig. 4.11, and the positive curvature still exists.

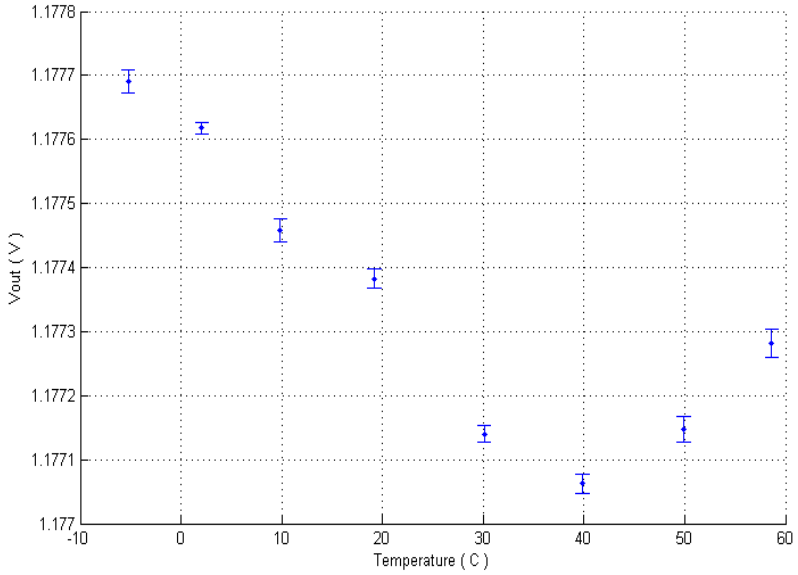

Figure 4.10 Measured bandgap voltage output transfer curve

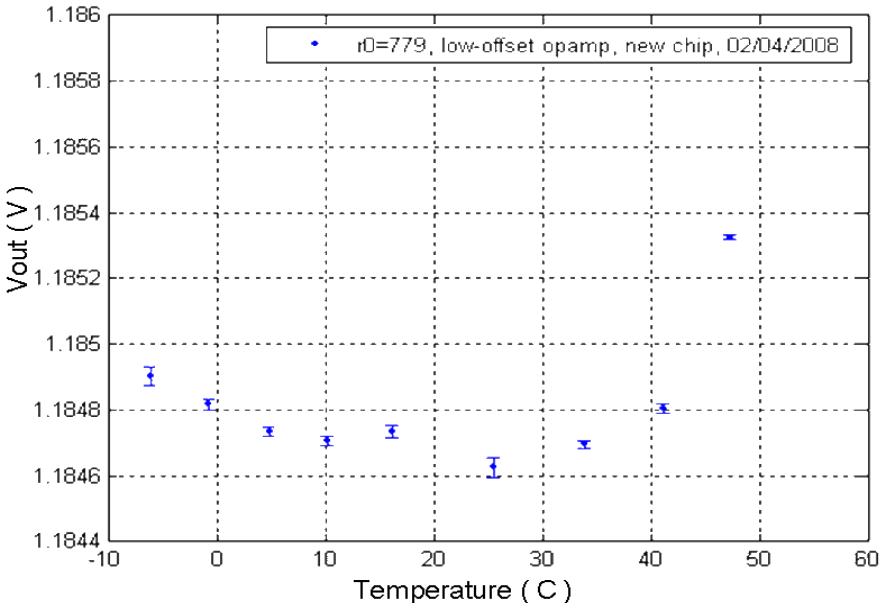

Figure 4.11 Measured output transfer curve using ultra low offset op amp

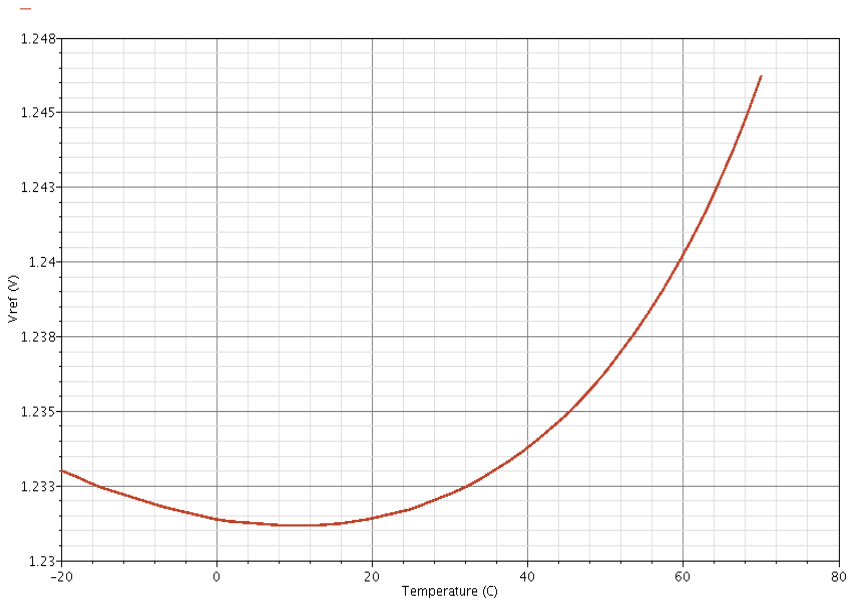

Figure 4.12 Positive curvature in the simulation when resistors have large second order TC

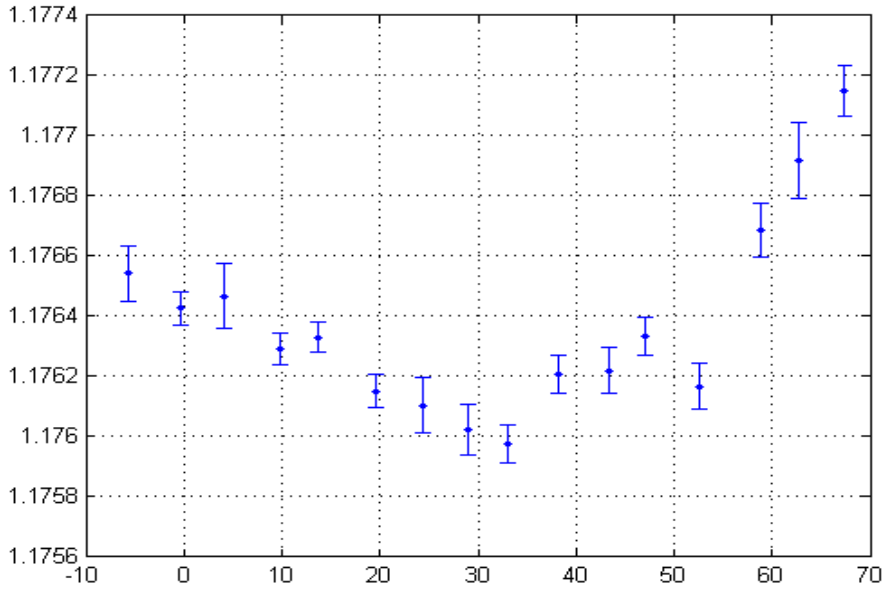

Figure 4.13 Transfer function with replaced diode chip II

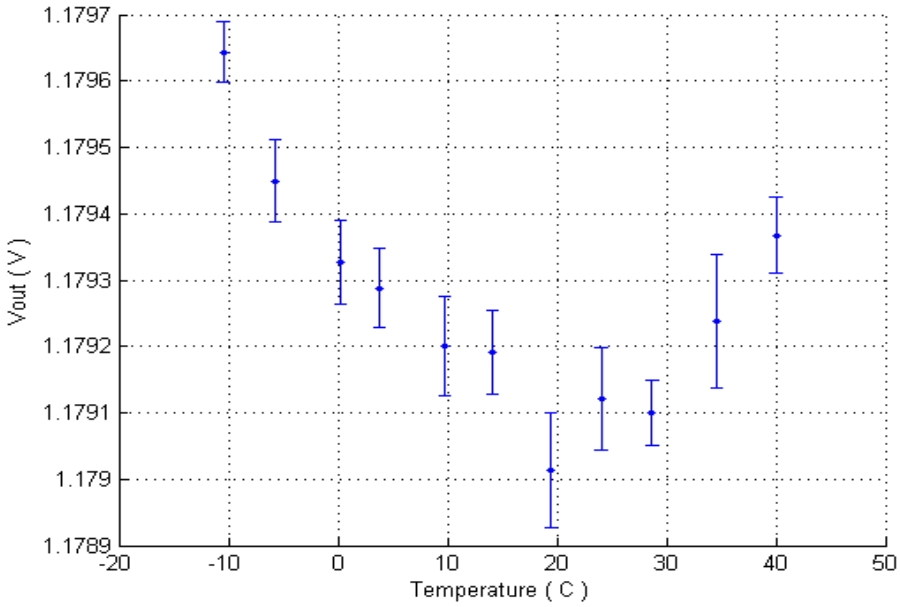

Figure 4.14 Transfer function with replaced diode chip III

Simulations in Spectre show that when the gain-determining resistor has a second-order temperature coefficient at the -1e-4 K^{-1} level, the second derivative will be positive as shown in Fig.4.12. I measured the temperature characteristics of the thin film resistors. From these measurements, I found the second-order temperature coefficient at the 0.086e-6 K^{-1} level, which is three orders of magnitude smaller than needed for the positive curvature.

To determine if there was something peculiar on the diode array, the chip containing the diodes (designated as chip I) was replaced with another two chips, designated as chip II and chip III. The measured results in Fig.4.13 and Fig. 4.14 for chip II and chip III, respectively, both have the positive curvature pattern.

The temperature coefficient TC from each measurement is at the level of 7~15 ppm/^{o}C, which corresponds to the 9~10-bit level. This thermal stability performance is at a level similar to or better than that reported in [2]–[10].

The cause of the positive curvature has not been determined but if predictable, could be useful in improving the thermal stability of voltage references. The effect of the nwell base spreading resistor associated with this particular on-chip diode layout pattern is one of the possible causes of the positive curvature but the effects of the base spreading resistance have not been quantified. The nwell spreading resistor is more non-linear than thin film resistors are. Future work focusing on identifying the cause of the change in sign of curvature in the measured results not apparent from computer simulations would be interesting. If the positive curvature can be predictably determined, it may be possible to improve the overall performance by combining the positive and negative curvature properties. If a circuit that combines a positive and a negative curvature effect could be designed, it is very likely that the relative weighting could be selected so that the resulting reference will have a near zero curvature, thus providing a level of thermal stability not achievable with existing approaches.

REFERENCES

[1] D. F. Hilbiber, “A new semiconductor voltage standard,” *Dig. Techn. Papers ISSCC*, pp.32-33. Feb. 1964.

[2] R. Widlar, “Designing Positive Voltage Regulators,” *EEE*, Vol. 17, pp. 90-97, June 1969.

[3] R. Widlar, “New Developments in IC Voltage Regularors”, *IEEE Journal of Solid State Circuits*, Vol. 6, pp. 2-7, Feb. 1971.

[4] P. Brokaw, “A Simple Three-Terminal IC Bandgap Reference”, *IEEE Journal of Solid State Circuits*, Vol. 9, pp.388-393, Dec.1974.

[5] G. Rincon-Mora, P. E. Allen, "A 1.1-V current-mode and piecewise-linear curvature-corrected bandgap reference," *IEEE Journal of Solid-State Circuits*, Vol. 33, pp. 1551 - 1554, Oct.1998.

[6] A. Buck, C. McDonald, S. Lewis and T.R. Viswanathan, "A CMOS bandgap reference without resistors," *IEEE Journal of Solid-State Circuits*, Vol. 37, No. 1, pp. 81-83, Jan. 2002.

[7] A. J. Annema, "Low-power bandgap references featuring DTMOSTs," *IEEE Journal of Solid-State Circuits*, Vol. 34, pp. 949 - 955, July 1999.

[8] H. Banba, H. Shiga, A. Umezawa, T. Miyaba, T. Tanzawa, S. Atsumi, K. Sakui, "A CMOS bandgap reference circuit with sub-1-V operation," *IEEE Journal of Solid-State Circuits*, Vol. 34, pp. 670 - 674, May 1999.

[9] B.-S. Song and P. R. Gray, "A precision curvature-compensated CMOS bandgap reference," *IEEE J. Solid-State Circuits*, Vol. SC-18, pp. 634–643, Dec. 1983.

[10] K. Kuijk, "A Precision Reference Voltage Source", *IEEE Journal of Solid State Circuits*, Vol. 8, pp. 222-226, June 1973.

[11] M. Gunawan, G.C.M. Meijer, J. Fonderie, and J.H. Huijsing, "A curvature-corrected low-voltage bandgap reference," *IEEE Journal of Solid State Circuits*, vol.28, no.6, pp.667-670, June 1993.

[12] Y. P. Tsividis, *Operation and Modeling of the MOS Transistor*. New York: McGraw-Hill, 1987.

[13] Linear Technology, "Zero Drift Operational Amplifer," Nov. 1999 http://cds.linear.com/docs/Datasheet/1052fa.pdf,

[14] Texas Intruments, “General Purpose Operational Amplifer,” Sep. 2000 http://www.datasheetcatalog.org/datasheet/texasinstruments/ua741.pdf

CHAPTER 5
A HIGHLY LINEAR VERY COMPACT UNTRIMMED ON-DIE TEMPERATURE SENSOR WITH SECOND AND THIRD-ORDER TEMPERATURE COMPENSATION

ABSTRACT

This paper proposes a family of CMOS structures as highly linear on-chip temperature sensors. As long as all transistors are in saturation, the output of the structure is a V_{DD} independent voltage that linearly expresses the CMOS threshold voltage, and hence is approximately linear in temperature. A sizing strategy is introduced following a combined analytical and numerical optimization approach, which effectively removes both second and third-order nonlinearities. With this sizing strategy, in this chapter and the following Chapter 6, three circuits that extract threshold voltage have been designed using 0.18 μm technology. Simulation results verify that the sensors' output voltages can be made very linear with temperature, with a simulated temperature INL (maximum temperature errors due to output voltage temperature nonlinearity) of around 0.05^{o}C over the temperature range of -20^{o}C ~100^{o}C. Results from both corners and Monte Carlo simulations demonstrate that the sensor linearity is robust over process variations and local device mismatch. With a standard two-point calibration, the sensor's maximum output error can be bounded by ± 0.15^{o}C without any trimming. The sensor is very compact, with a total layout area of around 400 μm^2 when implemented in the 0.18μm process.

I. INTRODUCTION

As component density continues to increase in advanced CMOS technologies, power density per unit die area of Very Large Scale Integration (VLSI) chips is increasing in many useful systems. Reliable operation of an integrated circuit system necessitates the prevention of excessive chip heating. Building on-chip temperature sensors to monitor the temperature at critical locations on a die is becoming an inevitable requirement. The on-chip measurement results also provide an opportunity to implement feedback from sensory data as a part of the power/thermal management algorithm. Due to the need for many sensors throughout the die, these on-chip sensors must be very compact. Since device reliability is highly sensitive and has a non-linear dependence on temperature, these temperature sensors must have measurement accuracy in the sub 1^{o}C range or better. Furthermore, to avoid self-heating, these sensors must have low power consumption.

The most widely researched temperature sensors are based upon the traditional proportional to absolute temperature (PTAT) principle and utilize the temperature-dependent characteristics of the pn junction to generate a PTAT voltage. Although this technique is widely used for building stand-alone temperature sensors, pnp elements and Op Amps are normally required to build those PTAT temperature sensors, which leads to larger die sizes and increased power consumption [1]–[4]. Other authors have focused on using the temperature dependence of the threshold voltage V_{th} and mobility of CMOS transistors to generate a temperature-dependent signal. The resultant signals have often been either a pulse width or an oscillation frequency that carries the temperature information. The reported circuits combine the effects of the temperature dependence of both mobility and threshold voltage, and are not highly linear with respect to temperature

[5]–[8]. Reported temperature errors with these approach range from +/-0.6^{o}C to several degrees Celsius. Recently, a temperature sensor that extracts CMOS threshold voltage has been reported to have inaccuracy at -1-0.8^{o}C and extremely small area around 50μm^2 [18]. However, the extraordinary small area only contains the part of the temperature sensor—the four NMOS transistors, while PMOS current mirror was realized by using external parameter analyzer.

In this work, a CMOS temperature sensor structure that is fully integrated, more compact and more linear with respect to temperature is introduced. In section II, a V_{DD} independent temperature sensor circuit that can express the threshold voltage is discussed. In section III, a combined analytical/numerical approach is used to "optimize" the sizes of the transistors in the temperature sensor structure to obtain good linearity with respect to temperature. In this work, the word "optimize" refers to a procedure that provides significant improvements in linearity from one step to the next in a sequence of operations with provisions for terminating the sequence of operations after a modest number of steps. Although the goal may be similar, the term "optimize" and the related concept of "optimization" usually have more stringent interpretations in the mathematics community. The objective of our work is to design a temperature sensor with very good temperature linearity, specifically, with linearity that is very good compared to that obtainable with the existing state of art designs [1] [3] [5]-[7] [14]-[18]. In the optimization presented in this work, the sequence of operations will be terminated either when the temperature linearity reaches the 0.05^{o}C level or when successive steps in the process no longer provide a predetermined reduction in linearity relative to the initial nonlinear error. From a mathematical point of view, the optimization obtained by this process may still not

provide an optimum solution or even a local minimum solution. However, from the circuit design point of view, the resultant temperature linearity of our design is very good when compared with other published results. Design insight on how to achieve trade-offs between linearity, area, and power consumption are also provided. A circuit design example is given in Section IV using the optimized sizing strategy introduced in Section III. The results are summarized in Section V.

II. THRESHOLD EXTRACTION CIRCUIT DESIGN

It is well known that the threshold voltage of a MOS transistor is highly linear with temperature. The relationship between threshold voltage and temperature will be discussed later in this section. Thus, a circuit that expresses the threshold voltage at the output can serve as a linear temperature sensor. In this section, emphasis will be placed on developing circuits that express the threshold voltage at the output.

A circuit that can express threshold voltage is shown in Fig.5.1. Although this circuit structure is not new, the use of this structure to generate a highly linear temperature-dependent output voltage is new. Temporarily neglecting the channel length modulation, four equations can be written to fully describe the operation of the circuit:

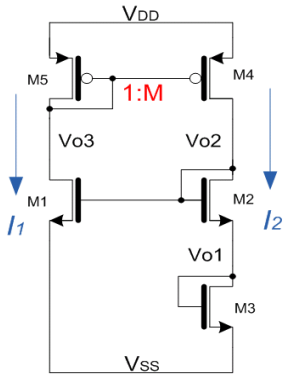

Figure 5.1 Schematic of the proposed temperature sensor circuit A

$$I_1 = \frac{\mu_n C_{ox}}{2}\frac{W_1}{L_1}(V_{o2} - V_{tn1})^2 \tag{5.1}$$

$$I_2 = \frac{\mu_n C_{ox}}{2}\cdot\frac{W_2}{L_2}(V_{o2} - V_{o1} - V_{tn2})^2 \tag{5.2}$$

$$I_2 = \frac{\mu_n C_{ox}}{2}\frac{W_3}{L_3}(V_{o1} - V_{tn3})^2 \tag{5.3}$$

$$I_2 = M \cdot I_1 \tag{5.4}$$

where V_{tn1}, V_{tn2}, V_{tn3} are the threshold voltages for M_1, M_2 and M_3 respectively, W and L variables denote the widths and lengths of the corresponding transistors, and M is the gain of the p-channel current mirror. Equations (5.1)–(5.4) comprise a set of four simultaneous equations in the unknown variables $\{I_{D1}, I_{D2}, V_{o1}$, and $V_{o2}\}$. V_{o1} and V_{o2} can be solved from these four equations to obtain

$$V_{o1} = \frac{(V_{tn1} - V_{tn2})\cdot\sqrt{\frac{W_2/L_2}{W_3/L_3}} + V_{tn3}\cdot\left(1 - \sqrt{\frac{W_2/L_2}{M\cdot(W_3/L_3)}}\right)}{1 + \sqrt{\frac{W_2/L_2}{W_3/L_3}} - \sqrt{\frac{W_2/L_2}{M\cdot(W_1/L_1)}}} \tag{5.5}$$

$$V_{o2} = \frac{V_{tn1}\cdot\left(1 + \sqrt{\frac{W_2/L_2}{W_3/L_3}}\right) - V_{tn2}\cdot\sqrt{\frac{W_2/L_2}{M\cdot(W_1/L_1)}} - V_{tn3}\cdot\sqrt{\frac{W_2/L_2}{M\cdot(W_1/L_1)}}}{1 + \sqrt{\frac{W_2/L_2}{W_3/L_3}} - \sqrt{\frac{W_2/L_2}{M\cdot(W_1/L_1)}}} \tag{5.6}$$

Assuming that all NMOS transistors have the same threshold voltage, (5.5) and (5.6) can be simplified to:

$$V_{o1} = V_{tn}\frac{1 - \sqrt{\frac{W_2/L_2}{M\cdot(W_1/L_1)}}}{1 + \sqrt{\frac{W_2/L_2}{W_3/L_3}} - \sqrt{\frac{W_2/L_2}{M\cdot(W_1/L_1)}}} \tag{5.7}$$

$$V_{o2} = V_{tn} \frac{1 + \sqrt{\frac{(W/L)_2}{(W/L)_3}} - 2\sqrt{\frac{(W/L)_2}{M \cdot (W/L)_1}}}{1 + \sqrt{\frac{(W/L)_2}{(W/L)_3}} - \sqrt{\frac{(W/L)_2}{M \cdot (W/L)_1}}} \tag{5.8}$$

From (5.7) and (5.8), it can be observed that the output voltages, V_{o1} and V_{o2}, will have a nearly linear relationship with the threshold voltage, V_{th}. According to the threshold voltage temperature dependence model in (5.9), the threshold voltage itself is nearly linear with respect to temperature [9]:

$$V_{tn}(T) = V_{tn0} + (KT1 + KT1L / L_{eff} + KT2 \cdot V_{bseff}) \cdot (T / T_{NOM} - 1) \tag{5.9}$$

where *KT1*, *KT1L*, and *KT2* are process dependent constants, T_{NOM} is equal to 300 K, L_{eff} is the effective length and is approximately equal to the length L, and V_{bseff} is the effective bulk to source voltage. From (5.9), it can be seen that if the CMOS bulk terminal is connected to source, the temperature nonlinearity brought by bulk to source voltage V_{bs} is negligible. In the circuit in Fig.5.1, a zero V_{bs} of M_1 and M_3 can be easily realized, while in M_2, the source and bulk cannot be easily tied together in most processes when a double-well is not available. This phenomenon suggests that the PMOS counterpart of circuit in Fig.5.1 will potentially have better temperature linearity because each PMOS device can have its own well tie.

III. SIZE STRATEGIES TO REDUCE TEMPERATURE NONLINEARITY

The channel modulation effect neglected in Section II will cause temperature nonlinearity in the output voltages V_{o1} and V_{o2} in the circuit of Fig.5.1. This type of nonlinearity will result in several degree Celsius temperature errors. To improve the

linearity, the channel modulation parameter, λ is re-introduced in the analytical model in (5.1)–(5.4) resulting in the more nonlinear circuit equations:

$$I_1 = \frac{\mu_p C_{ox}}{2}\frac{W_5}{L_5}(V_{DD} - V_{o3} - |V_{tp}|)^2[1 + \lambda_p(V_{DD} - V_{o3})] \tag{5.10}$$

$$I_1 = \frac{\mu_n C_{ox}}{2}\frac{W_1}{L_1}(V_{o2} - V_{tn1})^2(1 + \lambda_n V_{o3}) \tag{5.11}$$

$$I_2 = \frac{\mu_p C_{ox}}{2}\frac{W_4}{L_4}(V_{DD} - V_{o3} - |V_{tp}|)^2[1 + \lambda_n(V_{DD} - V_{o2})] \tag{5.12}$$

$$I_2 = \frac{\mu_n C_{ox}}{2}\frac{W_2}{L_2}(V_{o2} - V_{o1} - V_{tn2})^2[1 + \lambda_n(V_{o2} - V_{o1})] \tag{5.13}$$

$$I_2 = \frac{\mu_n C_{ox}}{2}\frac{W_3}{L_3}(V_{o1} - V_{tn3})^2(1 + \lambda_n \cdot V_{o1}) \tag{5.14}$$

By eliminating I_1 and I_2, the five equations (5.10)–(5.14) can be reduced to three equations (5.15)–(5.17) with independent variables $\{V_{o1}, V_{o2}, V_{o3}\}$.

$$\frac{\mu_p C_{ox}}{2}\frac{W_5}{L_5}(V_{DD} - V_{o3} - |V_{tp}|)^2[1 + \lambda_p(V_{DD} - V_{o3})] = \frac{\mu_n C_{ox}}{2}\frac{W_1}{L_1}(V_{o2} - V_{tn1})^2(1 + \lambda_n V_{o3}) \tag{5.15}$$

$$\frac{\mu_p C_{ox}}{2}\frac{W_4}{L_4}(V_{DD} - V_{o3} - |V_{tp}|)^2[1 + \lambda_n(V_{DD} - V_{o2})] =$$

$$\frac{\mu_n C_{ox}}{2}\frac{W_2}{L_2}(V_{o2} - V_{o1} - V_{tn2})^2[1 + \lambda_n(V_{o2} - V_{o1})] \tag{5.16}$$

$$\frac{\mu_n C_{ox}}{2}\frac{W_2}{L_2}(V_{o2} - V_{o1} - V_{tn2})^2[1 + \lambda_n(V_{o2} - V_{o1})] = \frac{\mu_n C_{ox}}{2}\frac{W_3}{L_3}(V_{o1} - V_{tn3})^2(1 + \lambda_n \cdot V_{o1})$$

(5.17)

By re-organizing (5.15)–(5.17), we can get:

$$f_1(V_{o1}, V_{o2}, V_{o3}, k_{11}...k_{1n}) = 0 \tag{5.18}$$

$$f_2(V_{o1}, V_{o2}, V_{o3}, k_{21}.....k_{2m}) = 0 \tag{5.19}$$

$$f_3(V_{o1}, V_{o2}, V_{o3}, k_{31}.....k_{3p}) = 0 \tag{5.20}$$

where V_{o1}, V_{o2} and V_{o3} are explicitly indicated and where the $k_{11}....k_{1n}$, $k_{21}.....k_{2m}$, $k_{31}....k_{3p}$ are not dependent upon these three port voltages and include the independent voltage source, V_{DD} and model parameters, such as V_{tp}, V_{tn1}, V_{tn2}, $\mu_n C_{ox} W_1/L_1$, $\mu_n C_{ox} W_2/L_2$, etc.

The highly non-linear form of equations (5.18)–(5.20) makes an explicit expression for V_{o1}, V_{o2} and V_{o3} difficult or impossible to obtain. As an alternative, we consider a power series expansion for the output voltages of interest, V_{o1} and V_{o2}, in terms of the temperature T, which can be expressed as:

$$V_{o1} = V_{o1n} + \frac{\partial V_{o1}}{\partial T}\cdot(T - T_0) + \frac{\partial^2 V_{o1}}{\partial T^2}\cdot(T - T_0)^2 + \frac{\partial^3 V_{o1}}{\partial T^3}\cdot(T - T_0)^3 + ... + \frac{\partial^n V_{o1}}{\partial T^n}\cdot(T - T_0)^n \tag{5.21}$$

$$V_{o2} = V_{o2n} + \frac{\partial V_{o2}}{\partial T}\cdot(T - T_0) + \frac{\partial^2 V_{o2}}{\partial T^2}\cdot(T - T_0)^2 + \frac{\partial^3 V_{o2}}{\partial T^3}\cdot(T - T_0)^3 + ... + \frac{\partial^n V_{o2}}{\partial T^n}\cdot(T - T_0)^n \tag{5.22}$$

In these equations, T_0 is the temperature about which the expansion is made. V_{o1n} and V_{o2n} are the nominal values of V_{o1} and V_{o2} at temperature T_0, and all derivatives are evaluated at T_0. Good linearity would be expected if the second-order partial derivatives and ideally the third-order partial derivatives are small or zero. The most straightforward way to find d^2V_{o1}/dT^2 and d^2V_{o2}/dT^2 is to solve V_{o1} and V_{o2} directly from (5.18)–(5.20). However, as observed previously, the highly non-linear forms of V_{o1} and V_{o2} make it very tedious to first solve for V_{o1} and V_{o2} directly and then apply differentiation.

Using the chain rule for implicit function differentiation, we can directly differentiate equations (5.18)–(5.20) with respect to temperature and then solve for the first, second and higher-order temperature derivative terms [10]. From the chain rule, we obtain the following equations:

$$\frac{\partial f_1}{\partial V_{o1}}\frac{\partial V_{o1}}{\partial T}+\frac{\partial f_1}{\partial V_{o2}}\frac{\partial V_{o2}}{\partial T}+\frac{\partial f_1}{\partial V_{o3}}\frac{\partial V_{o3}}{\partial T}+\sum_{i=1}^{n}\frac{\partial f_1}{\partial k_{1i}}\frac{\partial k_{1i}}{\partial T}=0 \tag{5.23}$$

$$\frac{\partial f_2}{\partial V_{o1}}\frac{\partial V_{o1}}{\partial T}+\frac{\partial f_2}{\partial V_{o2}}\frac{\partial V_{o2}}{\partial T}+\frac{\partial f_2}{\partial V_{o3}}\frac{\partial V_{o3}}{\partial T}+\sum_{i=1}^{m}\frac{\partial f_2}{\partial k_{2i}}\frac{\partial k_{2i}}{\partial T}=0 \tag{5.24}$$

$$\frac{\partial f_3}{\partial V_{o1}}\frac{\partial V_{o1}}{\partial T}+\frac{\partial f_3}{\partial V_{o2}}\frac{\partial V_{o2}}{\partial T}+\frac{\partial f_3}{\partial V_{o3}}\frac{\partial V_{o3}}{\partial T}+\sum_{i=1}^{p}\frac{\partial f_3}{\partial k_{3i}}\frac{\partial k_{3i}}{\partial T}=0 \tag{5.25}$$

From (5.23)-(5.25), it follows that equations (5.23)–(5.25) can be expressed as in (5.26).

$$\begin{bmatrix} 0 & g_{m1} & g_{m5} \\ g_{o2}+g_{m2} & -g_{o2}-g_{o4}-g_{m2} & -g_{m4} \\ g_{o2}+g_{o3}+g_{m2}+g_{m3} & -g_{o2}-g_{m2} & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial V_{o1}}{\partial T} \\ \frac{\partial V_{o2}}{\partial T} \\ \frac{\partial V_{o3}}{\partial T} \end{bmatrix} = \begin{bmatrix} K_1 \\ K_2 \\ K_3 \end{bmatrix} \tag{5.26}$$

In (5.26), the terms on the right side are not dependent upon the derivatives of the port voltage variables and for convenience, the small signal terms have been used to express the observed relationships between large signal model parameters and operating points. These small-signal terms are the standard trans-conductance g_m and output conductance g_o of the corresponding devices at the nominal operating point and at the expansion temperature T_0. The parameters, K_1, K_2, and K_3 are given by the expressions (5.27)–(5.29):

$$K_1 = g_{m1}\cdot\frac{\partial V_{tn1}}{\partial T}-g_{m5}\cdot\frac{\partial V_{tp}}{\partial T} \tag{5.27}$$

$$K_2 = g_{m4} \cdot \frac{\partial V_{tp}}{\partial T} - g_{m2} \cdot \frac{\partial V_{tn2}}{\partial T} \tag{5.28}$$

$$K_3 = -g_{m2} \cdot \frac{\partial V_{tn2}}{\partial T} + g_{m3} \cdot \frac{\partial V_{tn3}}{\partial T} \tag{5.29}$$

To simplify the derivation, when calculating dV_{o1}/dT and dV_{o2}/dT from (5.26), it is assumed that g_o is much smaller than g_m and can be neglected. Therefore, (5.26) can be simplified as:

$$\begin{bmatrix} 0 & g_{m1} & g_{m5} \\ g_{m2} & -g_{m2} & -g_{m4} \\ g_{m2}+g_{m3} & g_{m2} & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial V_{o1}}{\partial T} \\ \frac{\partial V_{o2}}{\partial T} \\ \frac{\partial V_{o3}}{\partial T} \end{bmatrix} = \begin{bmatrix} K_1 \\ K_2 \\ K_3 \end{bmatrix} \tag{5.30}$$

Although the simplification may introduce some errors, it significantly reduces the complexity of the derivation. Simulations of the overall temperature sensor discussed later show good correlation between the analytical models developed with these simplifications and the actual performance of the simulator.

Similarly, the second-order temperature derivative terms can be obtained by differentiating equations (5.23)–(5.25) with respect to temperature again. Following this approach, it follows that:

$$\begin{bmatrix} 0 & g_{m1} & g_{m5} \\ g_{o2}+g_{m2} & -g_{o2}-g_{o4}-g_{m2} & -g_{m4} \\ g_{o2}+g_{o3}+g_{m2}+g_{m3} & -g_{o2}-g_{m2} & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial^2 V_{o1}}{\partial T^2} \\ \frac{\partial^2 V_{o2}}{\partial T^2} \\ \frac{\partial^2 V_{o3}}{\partial T^2} \end{bmatrix} = \begin{bmatrix} K_{21} \\ K_{22} \\ K_{23} \end{bmatrix} \tag{5.31}$$

Again, by neglecting g_o, (5.31) can be simplified as follows:

$$\begin{bmatrix} 0 & g_{m1} & g_{m5} \\ g_{m2} & -g_{m2} & -g_{m4} \\ g_{m2}+g_{m3} & -g_{m2} & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial^2 V_{o1}}{\partial T^2} \\ \frac{\partial^2 V_{o2}}{\partial T^2} \\ \frac{\partial^2 V_{o3}}{\partial T^2} \end{bmatrix} = \begin{bmatrix} K_{21} \\ K_{22} \\ K_{23} \end{bmatrix} \tag{5.32}$$

where the terms K_{21}, K_{22}, and K_{23} do not contain any second derivatives of port voltage variables but may include the voltages V_{o1}, V_{o2}, and V_{o3} as well as the first-order derivatives that were previously determined.

Emphasis will now be placed on linearizing the voltage, V_{o1} with respect to temperature. If the assumption is made that the gain of the M_5:M_4 current mirror is 1, then $g_{m4}=g_{m5}$. It follows from (5.32) that the second order temperature derivative of V_{o1} can be explicitly expressed as:

$$d^2V_{o1} / dT^2 \approx \frac{K_{21}+K_{22}}{g_{m2}} \tag{5.33}$$

where

$$K_{21} = 2\lambda_p \cdot \mu_p C_{ox} \frac{W_5}{L_5} \cdot V_{EB5} \cdot (\frac{\partial V_{o3}}{\partial T} + \frac{\partial V_{tp}}{\partial T}) \cdot |\frac{\partial V_{o3}}{\partial T}| + \mu_p C_{ox} \frac{W_5}{L_5} (\frac{\partial V_{o3}}{\partial T} + \frac{\partial V_{tp}}{\partial T})^2$$
$$-2\lambda_n \mu_n C_{ox} \frac{W_1}{L_1} \cdot V_{EB1} (\frac{\partial V_{o2}}{\partial T} - \frac{\partial V_{tn1}}{\partial T}) \cdot \frac{\partial V_{o3}}{\partial T}$$
$$K_{22} = -2\lambda_p \cdot \mu_p C_{ox} \frac{W_4}{L_4} \cdot V_{EB4} \cdot (\frac{\partial V_{o3}}{\partial T} + \frac{\partial V_{tp}}{\partial T}) \cdot |\frac{\partial V_{o2}}{\partial T}| - \mu_p C_{ox} \frac{W_4}{L_4} (\frac{\partial V_{o3}}{\partial T} + \frac{\partial V_{tp}}{\partial T})^2$$
$$+2\lambda_n \mu_n C_{ox} \frac{W_2}{L_2} \cdot V_{EB2} (\frac{\partial V_{o2}}{\partial T} - \frac{\partial V_{o1}}{\partial T} - \frac{\partial V_{tn2}}{\partial T}) \cdot (\frac{\partial V_{o1}}{\partial T} - \frac{\partial V_{o2}}{\partial T})$$

To reduce the quadratic term expressed by (5.33), the objective is to make the sum of K_{21} and K_{22} close to zero. There are multiple solutions to realize this objective. One solution applied in this work is:

$$\frac{W_5}{L_5} = \frac{W_4}{L_4} \tag{5.34}$$

$$V_{EB4} \cdot \left| \frac{\partial V_{o3}}{\partial T} \right| = V_{EB5} \cdot \left| \frac{\partial V_{o2}}{\partial T} \right| \tag{5.35}$$

$$\frac{W_1}{L_1} \cdot V_{EB1} \left(\frac{\partial V_{o2}}{\partial T} - \frac{\partial V_{tn1}}{\partial T} \right) \cdot \frac{\partial V_{o3}}{\partial T} = \frac{W_2}{L_2} \cdot V_{EB2} \left(\frac{\partial V_{o2}}{\partial T} - \frac{\partial V_{o1}}{\partial T} - \frac{\partial V_{tn2}}{\partial T} \right) \cdot \left(\frac{\partial V_{o1}}{\partial T} - \frac{\partial V_{o2}}{\partial T} \right) \tag{5.36}$$

where dV_{o1}/dT, dV_{o2}/dT and dV_{o3}/dT in (5.35) and (5.36) are the values computed from (5.30).

These analytical results are first used to determine the size for each transistor and to reduce the second-order temperature non-linearities. This analytical approach will be followed by a numerical iteration, which further reduces the overall non-linearity. The analytical approach can be described as follows:

(a1) Select reasonable values, like two to five times the minimum length, for the lengths of M_4, M_5, M_2, M_3.

(a2) To satisfy conditions (5.34), the PMOS current mirror is chosen to have the same dimensions. Therefore, $L_4=L_5$, $W_4=W_5$.

(a3) It has been found that the size of M_1 is critical to the current consumption of the whole circuit. Consequently, the five variables $\{W_1, L_1, W_4 (W_5), W_2, W_3\}$ are chosen so that current consumption is within budget and the constraints given by (5.35) and (5.36) are satisfied as much as possible.

In this way, the quadratic term in V_{o1} can be significantly reduced. The temperature error caused by non-linearity can be maintained at about the 1^oC level by reducing the coefficient of the second-order term in (5.33). Ideally, the higher-order derivatives, such as d^3V_{o1}/dT^3 and d^4V_{o1}/dT^4 can be obtained in the same way as described

above. However, the tediousness of the resultant expressions grows rapidly when moving to a higher order.

To compensate for the third order temperature non-linear terms and to reduce temperature nonlinearity further, a finer size adjustment is required. A heuristic numerical sensitivity-based optimization procedure is used to reduce the temperature non-linearity. The design procedure can be described as follows:

(b1) Use the sizes obtained from the previous analytical approach as the initial sizes for the following numerical approach.

(b2) Keep the PMOS size fixed, because it has been found that PMOS size is less influential than that of the bottom three NMOS transistors on the output voltage temperature non-linearity. The sizes of W_1 and L_1 are determined from step (a3) so that the current consumption can be controlled. Therefore, two design variables $\{W_2, W_3\}$ are available for the size adjustment.

(b3) Vary W_2 by a certain amount ΔW_2, such as +20% or -20% of the original size, and find the sensitivity of the output temperature INL with respect to ΔW_2. If the INL decreases, replace the original value of W_2 with the new value. Otherwise, keep the original value of W_2.

(b4) Similarly, vary W_3 by the same percentage and find the sensitivity of the output temperature INL with respect to ΔW_3. If the INL decreases, replace the original value of W_3 with the new value. Otherwise, keep the original value of W_3.

(b5) Repeat step (b3) again and vary W_2 by a certain step amount that can be roughly determined according to the percentage change in the temperature INL when W_2 was varied by ΔW_2 in the previous iteration step.

(b6) Similarly, repeat step (b4) to vary W_3. The method for varying W_3 follows the same procedure as varying W_2 in step (b5).

(b7) Repeat the iterations in steps (b5)–(b6), and adjust W_2, W_3 until either of the two criteria is satisfied: (1) the percentage improvement of temperature INL relative to the original temperature INL at the very beginning of the optimization is less than 1%; (2) the obtained temperature INL is around 0.05^{o}C level. Any of the two criteria provides a termination criterion for the numerical iteration.

Then consider a finer adjustment in W_4 (W_5) and in the length of M_2 and M_3. This finer adjustment involves: first, perform iterations in changing W_4 (W_5) until the temperature INL does not have obvious improvement; second, adjust L_2 and L_3 in turns until the temperature INL is very small.

By using this combined analytical and numerical approach described above, the temperature INL can be reduced to a level that is less than 0.1^{o}C level.

IV. STARTUP CIRCUIT

Most reported MOS-based temperature sensors use some variant of a bias generator to provide a temperature-dependent output variable. These V_{DD}-independent bias generators typically have one or more feedback loops which invariably result in multiple stable equilibrium operating points. It is well-known that startup circuits must be added to ensure that the circuit will enter the desired operating state. However, in some circuits, the need for a start-up circuit may not be readily recognized and the methods often used to validate the performance of a startup circuit may not be effective. Also, the effectiveness of a start-up circuit may be difficult to observe in simulations or with analytical calculations. For example, repeated transient simulations of a circuit along

with a start-up throughout the design process and over corners may show that the circuit always starts up correctly, but after fabrication, the circuit may fail to start up correctly. If transient simulations of a circuit show that the circuit fails to start up and the same transient simulations show that the circuit always starts up when a start-up circuit is added, the designer may be tempted to conclude that the correct start-up is due to an effective start-up circuit. However, transient simulations provide only a single solution to a circuit and cannot guarantee that multiple stable equilibrium points are eliminated.

To circumvent this potential problem, a method for guaranteeing proper start-up is essential. An investigation of the circuit of Fig. 5.1 is useful for demonstrating how that can be done. One way to visualize the operation of this circuit is to observe that transistors M_1 and M_5 form a simple inverting amplifier with an input on the gate of M_1 and an output on the drain of M_5. Likewise, transistors M_2, M_3, and M_4 form a simple inverting amplifier with an input on the gate of M_4 and an output on the drain of M_2. These two inverters are connected in a two-inverter loop with no loading at DC. By breaking the loop at the high impedance node V_{g1} at the gate of M_1 as shown in Fig.5.2, a two-inverter loop is obtained. The transfer characteristics of this two-inverter loop are shown in Curve I in Fig. 5.3. Note that the transfer characteristics of this inverter intersect the $V_{out}=V_{in}$ line exactly three times. The upper and lower intersection points with slope K less than 1 are two stable equilibrium points and thus a start-up circuit is needed to guarantee it starts up at the desired equilibrium point. After a start-up circuit is added, the transfer characteristics of the loop are modified to Curve II shown in Fig.5.3. Since there is now a single equilibrium point, the circuit will start up at the desired operating point.

When designing the temperature sensor circuit using the topology in Fig.5.1, a simple start-up circuit comprised of a single transistor M_{st} is added as shown in Fig.5.4. Applying the method described above to break the loop and examining the intersection points between open loop transfer function and line $V_{out}=V_{in}$ will reveal all equilibrium points. The effective startup circuit should eliminate all solutions except the desired stable equilibrium point.

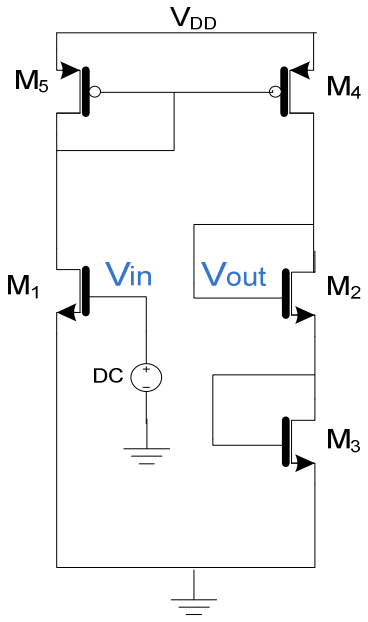

Figure 5.2 Conceptual circuit verifying the existence of multiple stable equilibrium points

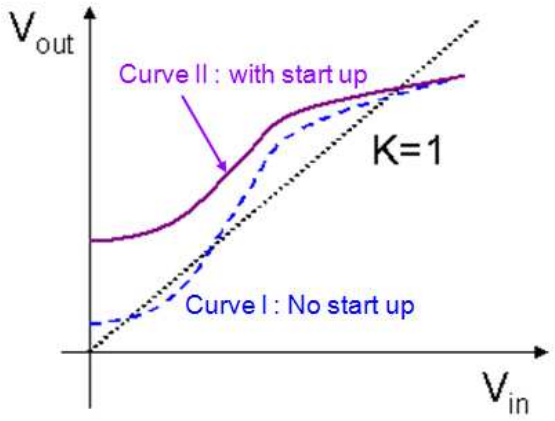

Figure 5.3 Illustration of existence of multiple stable equilibrium points and their removal by adding a start-up circuit

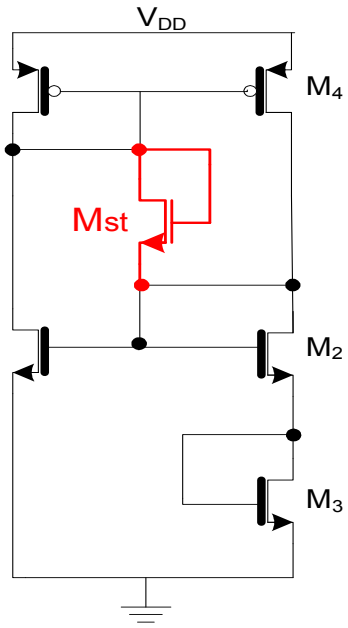

Figure 5.4 Illustration of a simple start-up circuit

V. DESIGN EXAMPLE AND ITS PERFORMANCES

To demonstrate the good temperature linearity property of the circuit in Fig.5.1 and the effectiveness of the sizing strategy, a circuit has been designed in a 1P6M 0.18μm process using the BSIM3v3 model. Define the temperature error caused by temperature non-linearity of the sensor's output voltage as the interpreted temperature

error from the difference between the transfer curve, V_{out} versus temperature and its ending point fit line. The temperature sensor's INL is then the maximum absolute value of the temperature error over the temperature range determined by the two ending-points fit line. After applying the sizing strategy described earlier to the sensor circuit in Fig. 5.1, the temperature error of the sensor at typical process conditions is shown in Fig.5.5. Note that the temperature error curve has three extremes and four zero-crossings in the temperature range of interests. Also, notice the nearly symmetric shape of the curve. Using our proposed design and sizing strategy, the temperature INL error can be reduced significantly by reducing the second and third-order temperature non-linearities in the sense of inner products [11], which means the circuit has been sized so that (5.37) and (5.38) are satisfied as close as possible.

$$< (V_{out}(T) - V_{fit}), P_2 > \approx 0 \tag{5.37}$$

$$< (V_{out}(T) - V_{fit}), P_3 > \approx 0 \tag{5.38}$$

where $V_{out}\ (T)$ is transfer curve of output voltage, V_{fit} is the best fitting line. P_2 and P_3 are basis function obtained from Gram-Schmidt orthogonalization process [12] [13]. They can be found from expressed in (5.39)-(5.41)

$$P_0 = 1 \tag{5.39}$$

$$P_2 = T - T_{mid} \tag{5.40}$$

$$P_k = [(T - T_{mid}) - \frac{< (T - T_{mid}) \cdot P_{k-1}, P_{k-1} >}{< P_{k-1}, P_{k-1} >}] \cdot P_{k-1} - \frac{< P_{k-1}, P_{k-1} >}{< P_{k-2}, P_{k-2} >} \cdot P_{k-2} \qquad \text{for k=2,3,4....} \tag{5.41}$$

where T_{mid} is the middle temperature of the overall temperature range from T_{min} to T_{max}.

It can be proved that in the temperature sensor design to satisfy (5.37) is equivalent to have d^2V_{o1}/dT^2 in (5.33) minimized. Therefore, following the analytical optimization procedure in section III can move the circuit design spaces near an optimized value for temperature INL error. In the neighborhood of the optimized result, the numerical optimization procedure as described in section III is performed so that (5.38) can be satisfied as close as possible, too. The temperature error magnitude is reduced dramatically relative to the temperature INL error curve before size optimization in Fig. 5.6 and the fourth-order nonlinearity effect becomes quite visible as shown in Fig.5.5. The temperature nonlinearity can be maintained at around 0.05^oC or less throughout the temperature range -20^oC to 100^oC.

The effect of global parameter variations from different corners is also investigated. Simulation results in Fig.5.7 show that the circuit has a worst-case temperature INL of 0.15^oC, which demonstrates good robustness of the design at different process corners. In the worst corner—slow NMOS slow PMOS, the voltage V_{o2} is higher due to the larger threshold voltage. The transistor M_4 loses headroom in its V_{DS} voltage, and therefore tends to operate near triode region and degrades the temperature linearity predicted in Section II.

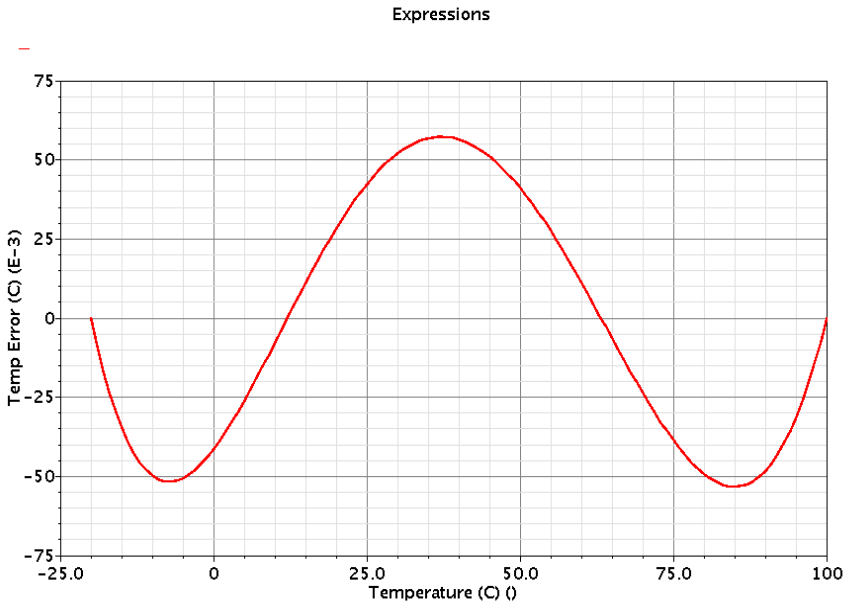

Figure 5.5 Temperature error of the proposed temperature sensor at the typical conditions

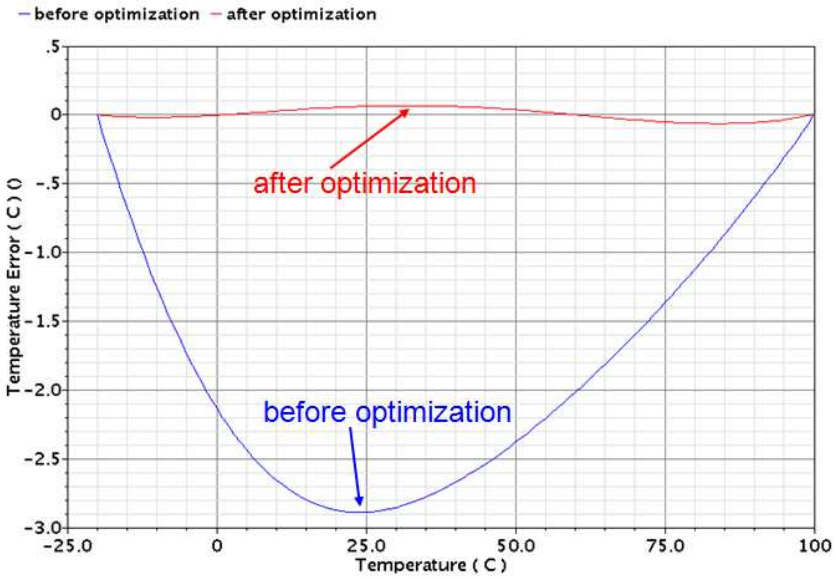

Figure 5.6 Temperature error curves before and after size optimization

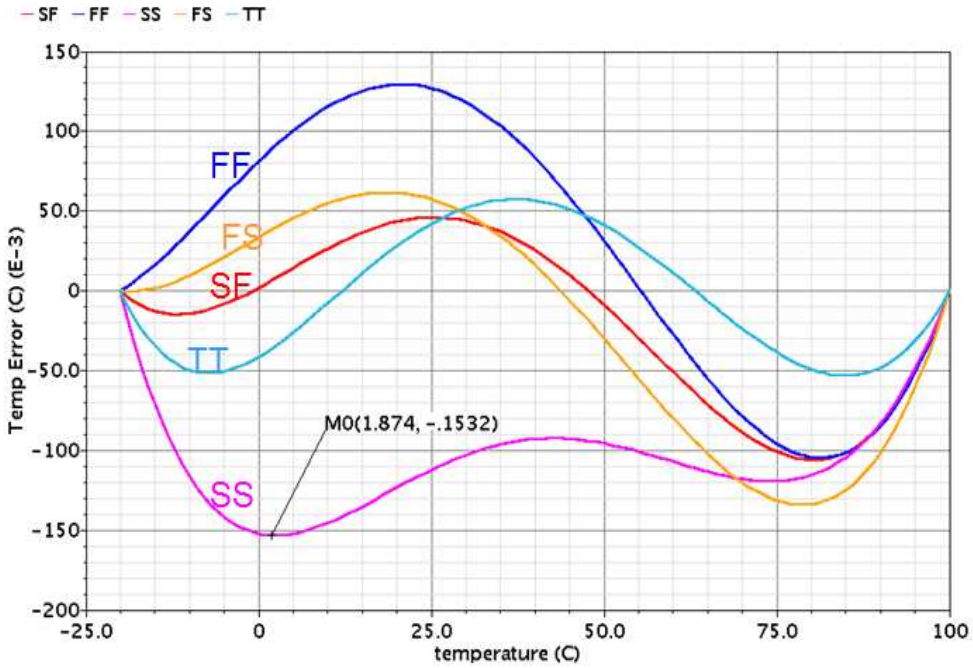

Figure 5.7 Temperature error curves at different process corners (TT: typical; FF: fast NMOS fast PMOS; FS: fast NMOS slow PMOS; SS: slow NMOS slow PMOS; SF: slow NMOS fast PMOS)

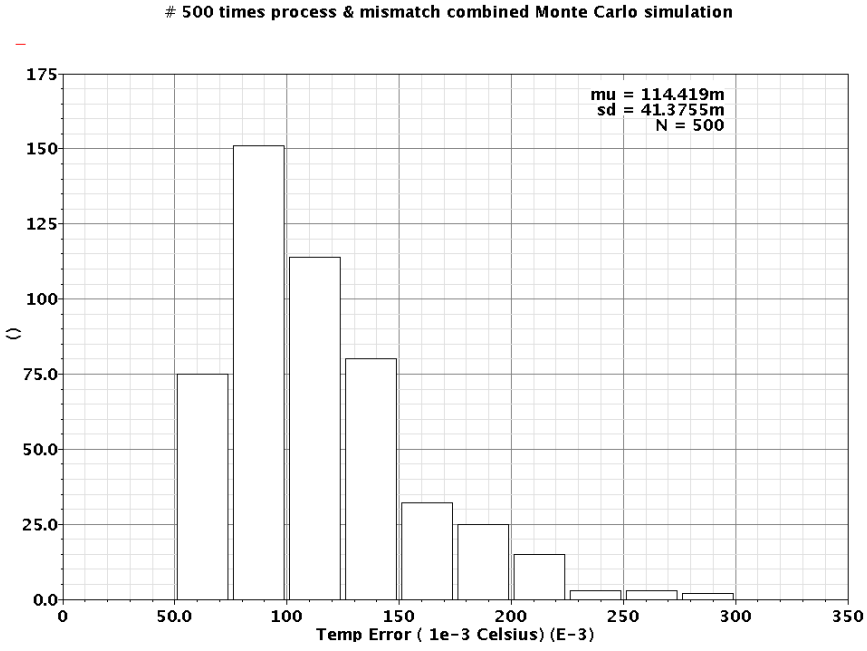

Figure 5.8 Distribution of worst case temperature error under process and mismatch combined Monte Carlo simulation

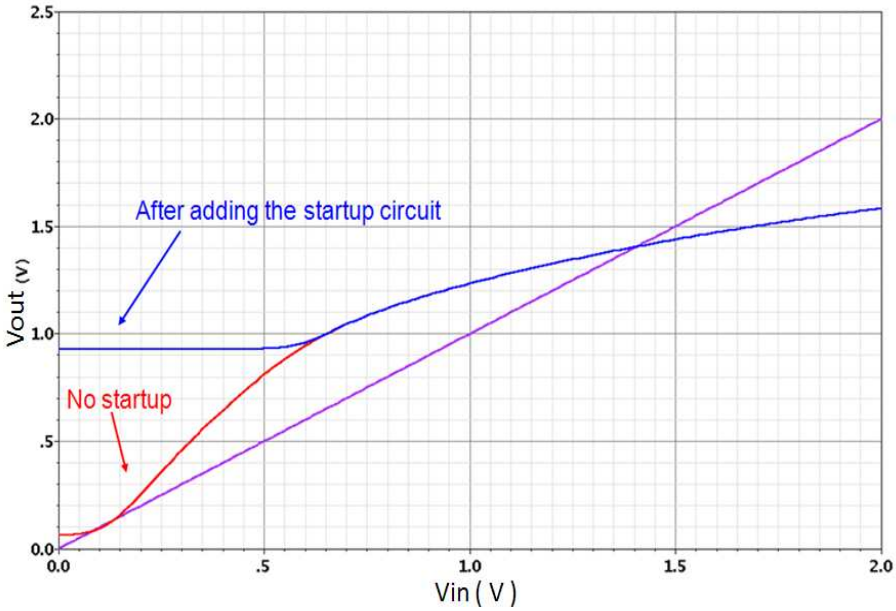

Figure 5.9 Start-up circuit avoids undesired operating points

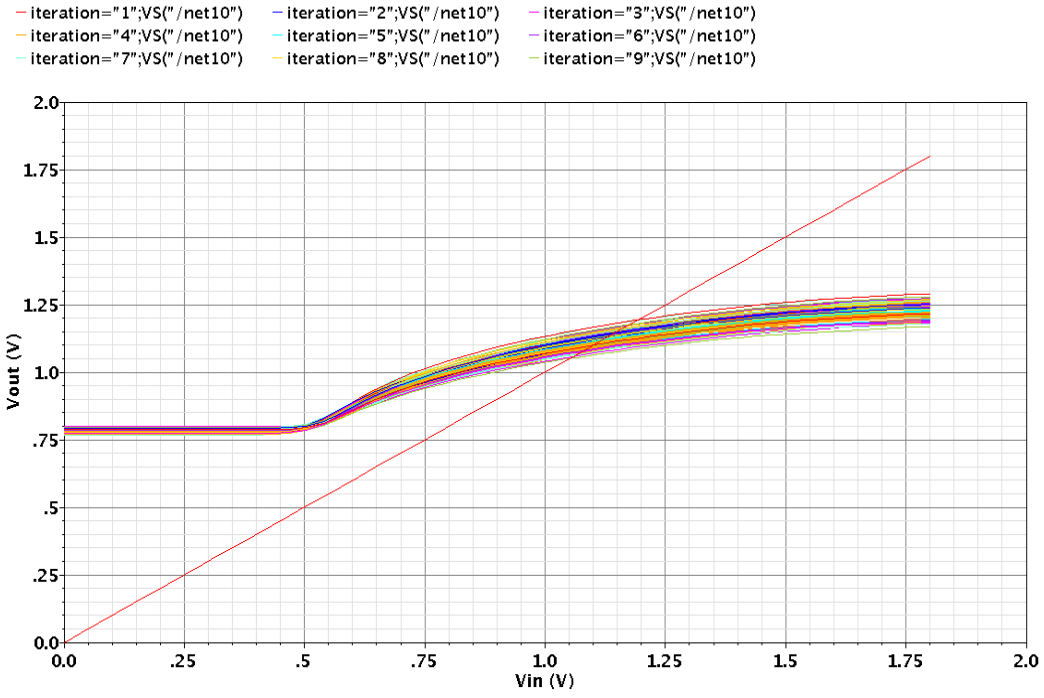

Figure 5.10 With process variations and device mismatches at T=27°C, the start-up circuit always avoids undesired operating points

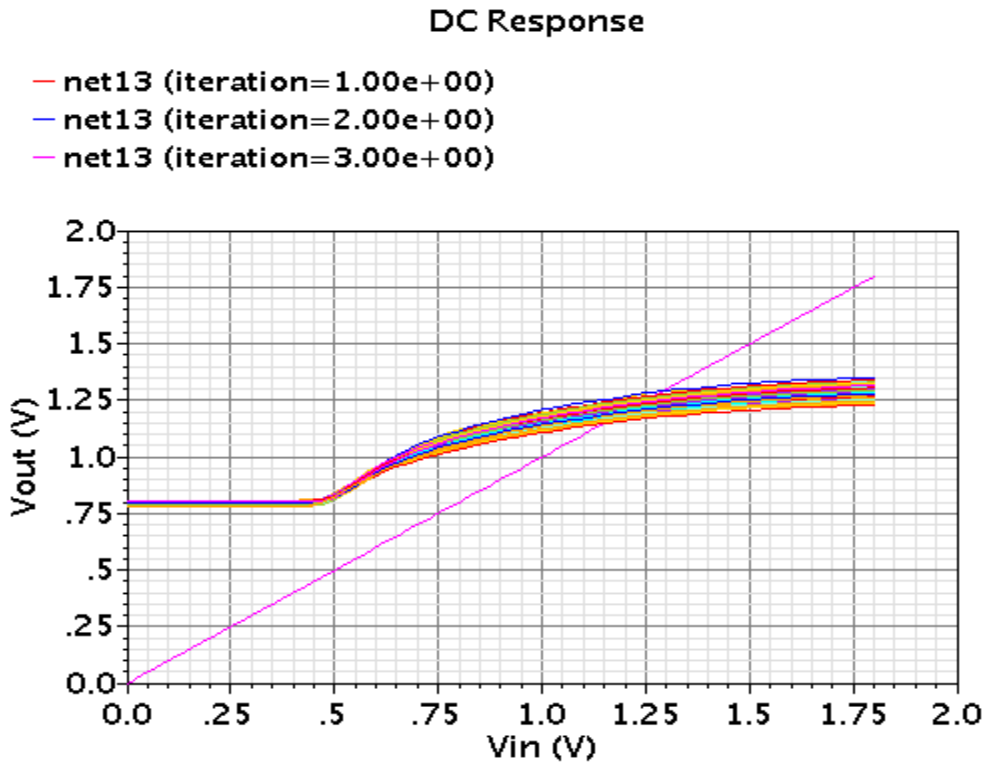

Figure 5.11 With process variations and device mismatches at T= -20°C, the start-up circuit always avoids undesired operating points

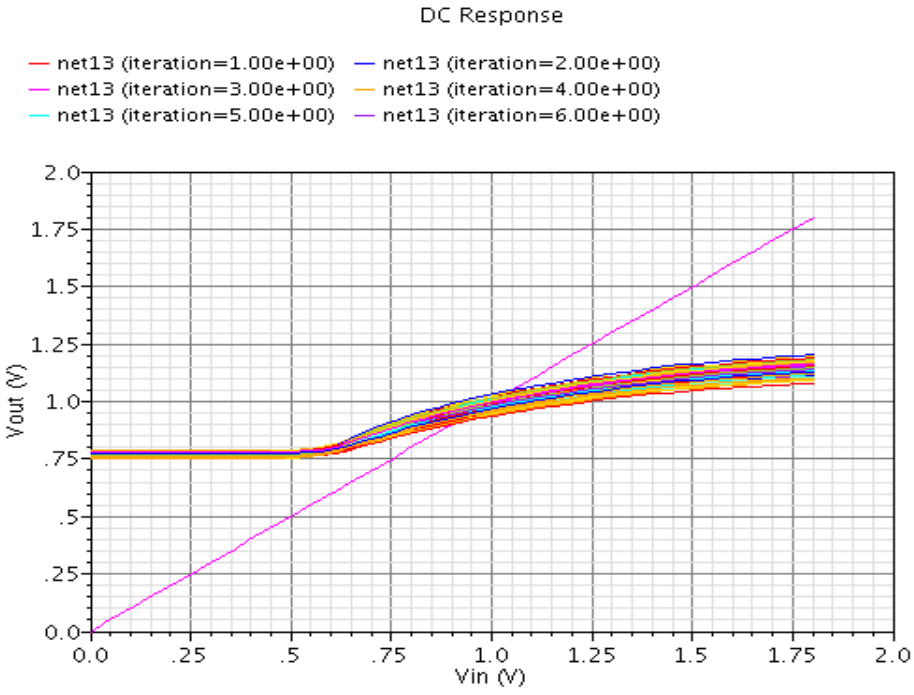

Figure 5.12 With process variations and device mismatches at T= 100°C, the start-up circuit always avoids undesired operating points

Table 5.1 Summary of performances

Parameters	*Performances*
Maximum Temp Error (°C) at typical condition	0.055
Maximum Temp Error (°C) at worst corner	0.16
Maximum Temp Error (°C) at 500 # MC simulation	mu=0.11, std=0.04
Current (μA)	53.5
Layout area estimated (μm^2)	400
Output Voltage Temp Coefficient (mV/°C)	-0.89

Table 5.2 Key circuit parameters

Process (μm)	0.18
V_{DD} (V)	1.8
Temperature range (°C)	-20 ~100
Transistor sizes	$(W/L)_1$=0.3 μ / 0.8 μ, $(W/L)_2$=2x10 μ / 0.38 μ, $(W/L)_3$=2x3.3μ/0.4 μ, $(W/L)_{st}$=6x0.5 μ / 0.72 μ, $(W/L)_4$=5 μ / 0.9 μ, $(W/L)_5$=5 μ / 0.9 μ

Table 5.3 Comparison with recent integrated temperature sensors

Sensors	**Max Temp Error (°C)**	**Area (mm^2)**	**Temperature Range (°C)**	**Power Consumption**	**CMOS technology**
PN junction based [1]	±1	~1.3	-40~120	7 μW	2 μm
PN junction based [3]	0.1 (one corner, 24 samples)	4.5	-55~125	> 187.5 μW	0.7μm
PN junction based[15] For PowerPC microprocessor	±4 with trim	N/A	10~100	N/A	N/A
Thermal diode, Intel [16]	~±8	N/A	40~110	N/A	N/A
Thermal diode, AMD [17]	~±10	N/A	-55~85	10~100 μW	90nm
TDC [5]	-0.7~+0.9	0.175	0~100	10μW	0.35μm
Current-freq converter [7]	~ 1	0.003~0.02	10~100	200 μW	1μm
Ring Oscillator based [6]	~7 (corners)	N/A (14 transistors)	40~150	25μW	45nm
Ring oscillator based [14]	+2.75 ~ -2.9	0.0013	-40~110	400μW	65nm
CMOS threshold voltage based [18]	-1 ~ 0.8	50e-6 Needs External Current Mirror	50~125	25μW	90nm
Proposed (circuit A)	0.055(TT) 0.16 (5 corners)	0.0004	-20~100	90μW	0.18μm

In Fig.5.8, results from 500 runs of Monte Carlo simulation that were run to test for not only process variations, but also device mismatch in the 0.18μ process using a BSIM3v3 model show that the circuit also has very good robustness when process variations and local device mismatch are both present. The Monte Carlo statistical models are provided as part of the model files for the process and include for both local and process variations in the threshold voltage, mobility, transistor width and length and possibly some other parameters. From these Monte Carlo simulations, it is observed that the mean value for the maximum temperature INL error is 0.11^{o}C, and the standard deviation is 0.04^{o}C. The main performances and design specifications have been listed in Table 5.1. The key circuit design parameters are listed in Table 5.2. The stable equilibrium points were verified by the loop transfer simulation before and after adding the start-up circuit at typical conditions as shown in Fig. 5.9. From Fig.5.9, it can be observed that with typical conditions, the added startup circuit is effective at eliminating other undesired operation points. To make sure the circuit can always start up with process variations and device mismatch, 200 runs of Monte Carlo simulation are made at T=27^{o}C, -20^{o}C and 120^{o}C, respectively and the results are shown in Fig.5.10, Fig.5.11 and Fig.5.12, respectively. These simulations show only one intersection point with the line V_{out}=V_{in} for each V_{out} versus temperature curve and the open loop transfer function indicates all will start up correctly.

The performances of the proposed temperature sensor are compared with the published temperature sensors' performances in Table 5.3 and demonstrates ultra small area occupation and highly linear with temperature. The results presented in this design are all based upon computer simulations. The models used in the simulation are the same

models used in industry today for this process. The highly linear temperature performance of this circuit is strongly dependent on the temperature dependent model of the threshold voltage given in (5.9). Experimental verification of the performance of this temperature sensor is important to determine if the temperature dependence of the MOS transistor is adequately captured in (5.9).

VI. CONCLUSIONS

In this paper, a compact on-chip temperature sensor has been proposed. This structure can express the threshold voltage of CMOS transistors as outputs and achieve high temperature linearity. A sizing strategy using a combined analytical and numerical approach has been described that significantly reduces second and third order temperature nonlinearity. The designed circuit demonstrates a temperature INL error at the 0.05^{o}C level and robustness to process variations and local device mismatches and power supply variations. The small area and high linearity make the structure suitable for high-precision, multi-site on-chip temperature measurements.

REFERENCES

[1] A. Bakker and J. Huijsing, "Micropower CMOS Temperature Sensor with Digital Output", *IEEE J. Solid State Circuits*, Vol. 2, pp. 933-937, July 1996.

[2] A. Bakker,"CMOS Smart Temperature Sensors – An Overview", *Proc. IEEE Sensors Conference*, pp 1423-1427, Vol. 2, 2002.

[3] M. A. P. Pertijs, K. A. A. Makinwa, and J. H. Huijsing, "A CMOS Smart Temperature Sensor with a 3σ Inaccuracy of $\pm 1^{o}$C from -55^{o}C to 125^{o}C", *IEEE J.Solid State Circuits*, Vo1. 40, pp. 2805-2815, Dec. 2005.

[4] M. Tuthill, "A switched-current, switched-capacitor temperature sensor in 0.6-μm CMOS," *IEEE J. Solid-State Circuits*, Vol. 33, no. 7, pp. 1117–1122, Jul. 1998.

[5] P. Chen, C. Chen, C.Tsa, and W. Lu, "A Time-to-Digital-Based CMOS Smart Temperature Sensor", *IEEE Journal of Solid State Circuits*, pp. 1642-1648, August 2005.

[6] B. Datta and W. Burleson, "Low-Power and Robust On-Chip Thermal Sensing Using Differential Ring Oscillators", *IEEE Midwest Symposium on Circuits and Systems*, pp. 29-32, August 2007.

[7] V. Szekely, C. Marta, Z. Kohari and M. Rencz, "CMOS Sensors for On-Line Thermal Monitoring of VLSI Circuits", *IEEE Trans. On VLSI Systems*, Vol. 5, pp. 270-276, September 1997.

[8] K. Arabi and B. Kaminska, "Built-In Temperature Sensors for On-Line Thermal Monitoring of Microelectronic Structures", *Proc. IEEE International Conference on Computer Design (ICCD)*, pp. 262-465, October 1997.

[9] X. Xi, M. Dunga, J. He, W.Liu and C. Hu, *BSIM4.3.0 MOSFET Model Users' Manual*. Berkeley, CA: Univ. California, pp. 12-1, 2001.

[10] W.Fulks, *Advanced Calculus: An introduction to analysis*. John Wiley& Sons, third edition, pp.321-327, 1978.

[11] D. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.

[12] G. Golub and C. F. Loan. *Matrix Computations (3rd edn)*. The John Hopkins University Press: Baltimore, 1996.

[13] Ake Bjorck. *Numerical Methods for Least Squares Problems*. SIAM: Philadelphia, PA, 1996.

[14] K. Kim, H. Lee, S. Jung, C. Kim, "366kS/s 400uW 0.0013mm2 Frequency-to-Digital Converter Based CMOS Temperature Sensor Utilizing Multiphase Clock," *IEEE Custom Intergrated Circuits Conference*, pp. 203-206, 2009.

[15] H. Sanchez, et al., "Thermal Management Systems for High Performance Power PC Microprocessors", *Proc. IEEE Compcon*, pp. 325-330, Feb. 1997.

[16] Intel Corporation, "*Mobile Intel 965 Express Chipset Family Datasheet, Revision 3*", June 2007.

[17] Advanced Micro Devices Inc, "*AMD Functional Data Sheet, 940 Pin Package, Rev 3.05*", June 2004

[18] M. Sasaki, M. Ikeda, M. Asada, "A temperature sensor with an in- accuracy of -1/+0.8◦C using 90-nm 1-V CMOS for online thermal monitoring of VLSI circuits", IEEE Trans. on Semiconductor Manufacturing, Vo1. 21, pp.201-207, May 2008.

CHAPTER 6
HIGHLY LINEAR VERY COMPACT UNTRIMMED ON-DIE TEMPERATURE SENSORS WITH IMPROVED IMMUNITY TO LOW VOLTAGE HIGH THRESHOLD VOLTAGE SCENARIOS

ABSTRACT

The on-die temperature sensor in Chapter 5 that extracts threshold voltage demonstrates excellent temperature linearity over process corners and device mismatch. However, its performance will be degraded at low V_{DD} condition and this degradation is significantly worse at high threshold corners. In this chapter, the reason for this temperature linearity degradation is carefully examined, and a simple approach is proposed to reduce the sensitivity of the temperature linearity to the power supply deduction and threshold increase. The approach only requires small modifications on the circuit in Fig.5.1. Based upon two design examples, after the modification, the maximum temperature error is reduced from 1.5^{o}C to less than 0.3^{o}C at the worst corner.

I. INTRODUCTION

Previous work in Chapter 5 proposed a CMOS based temperature sensor whose compact size and high linearity make it well suited for emerging on-chip temperature measurement applications. The sensor output voltage can correctly express the MOS threshold voltage, which is highly linear with temperature. The combined analytical and numerical approach to sizing optimization can effectively reduce the second and third-order temperature non-linearities, resulting in a typical temperature INL error of about 0.05^{o}C, which is about ten times better than the current state of the art in CMOS based

on-die temperature sensors [2]-[5]. The circuit is simulated under various process corners, device mismatches and power supply variations. The worst temperature linearity performance occurs at low voltage, slow NMOS slow PMOS corner. In that corner, the node voltage V_{o2} in Fig.5.1 that expresses the threshold voltage will increase by a significant larger amount due to the high threshold voltage at that corner. Meanwhile, the decrease of V_{DD} will squeeze the V_{DS} voltage of M_4 and drive M_4 into near triode region. Consequently, the temperature linearity will degrade.

From the above observations, any method to alleviate this problem must prevent excessive reduction in V_{DS}. Normally, since V_{DD} is provided by an external source, we do not have much control over that. The only possible way is to soften the increase of V_{o2} at the slow corners. In this chapter, we will introduce a simple approach to implement this idea with a very small amount of increased hardware.

II. LOW VOLTAGE HIGH THRESHOLD SCENARIO

The two circuits B and C will have better performances while operating at low V_{DD} and large threshold voltage V_{th} corners. In circuit B in Fig. 6.1, M_2 and M_6 form an active attenuator when M_6 is in the ohmic region and M_2 is in saturation [1]. The equivalent resistor of M_6 can be written as:

$$R_{eq} = \frac{1}{\mu C_{ox}(W_6 / L_6)\cdot(V_{o2} - V_{tn6})} \tag{6.1}$$

Since V_{o2} expresses the threshold voltage, the equivalent resistor of M_6 is proportional to the CMOS transistor threshold voltage. Therefore, R_{eq} will increase at high threshold corners. Since the equivalent resistor from M_6 increases at large V_{th} corners and current level will decrease, the excess bias voltages of M_2 and M_3 decrease.

V_{o2} in the three circuits A, B and C are equal to $V_{EB2}+V_{th2}+V_{EB3}+V_{th3}$, but V_{o2} in the circuits in Fig.6.1, will not increase as much as that in circuit A because the extra transistor M_6 reduces V_{EB2} and V_{EB3} by reducing the current. The benefit of using M_6 instead of directly adding a resistor to reduce the total current is that: after the modification, the whole circuit is still made of ALL CMOS transistors, which is critical in suppressing the non-linear temperature effect of mobility on the expressions for V_{o1} and V_{o2}. Besides, CMOS transistors will be able to track each other over process corners better.

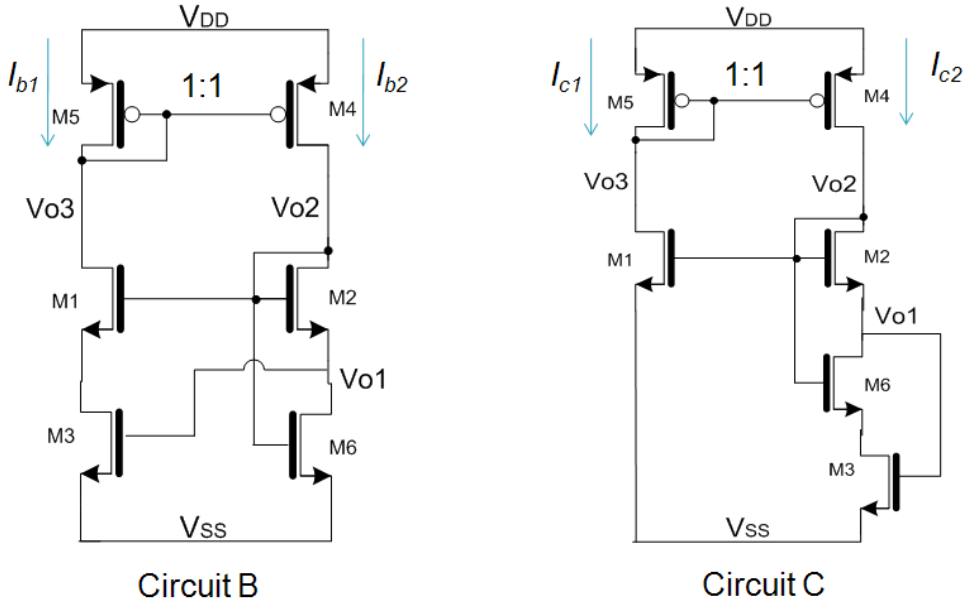

Figure 6.1 The proposed circuit B and circuit C with improved low voltage robustness

III. THRESHOLD EXTRACTION CIRCUITS WITH IMPROVED PERFORMANCES IN LOW V_{DD} HIGH THRESHOLD CORNERS

Circuit B and C are both threshold extraction circuits. Compared with circuit A in Fig.5.1, in the circuits in the Fig.6.1, the previous M_2 devices have been replaced by an

active attenuator. The tail diode connected device M_3 can be placed at either the left hand side or the right hand side. In circuit B, assuming the λ effect is neglected:

$$V_{o1} = \theta(V_{o2} - V_{tn2}) \tag{6.2}$$

where the attenuation factor θ is given by:

$$\theta = 1 - \sqrt{\frac{(W/L)_6}{(W/L)_6 + (W/L)_2}} \tag{6.3}$$

A set of equations can be written to describe the operation of circuit B in Fig.6.1:

$$I_{b1} = 1/2 \cdot \mu_n C_{ox} (W/L)_3 (V_{o1} - V_{tn3})^2 \tag{6.4}$$

$$I_{b2} = 1/2 \cdot \mu_n C_{ox} \cdot (W/L)_2 \cdot (V_{o2} - V_{o1} - V_{tn2})^2 \tag{6.5}$$

$$I_{b2} = I_{b1} \tag{6.6}$$

Combining (6.2)-(6.6), V_{o1} and V_{o2} in circuit B can be solved as:

$$V_{o1} = \frac{1}{1 - \sqrt{\frac{(W/L)_2}{(W/L)_3}}(\frac{1}{\theta} - 1)} \cdot V_{tn3} \tag{6.7}$$

$$V_{o2} = \frac{1}{\theta - \sqrt{\frac{(W/L)_2}{(W/L)_3}}(1-\theta)} \cdot V_{tn3} + V_{tn2} \tag{6.8}$$

Similarly, a set of equations (6.9)-(6.10) can be written to describe circuit C:

$$V_{o1} - V_{d3} = \theta \cdot (V_{02} - V_{tn2} - V_{d3}) \tag{6.9}$$

where V_{d3} is the drain voltage of M_3.

By sizing the devices M_2 and M_6 such that W_2=W_6. Then we can get:

$$I_{c2} = \frac{\mu_n C_{ox}}{2} \cdot \frac{W_2}{L_2 + L_6} \cdot (V_{o2} - V_{d3} - V_{tn2})^2 \tag{6.10}$$

$$I_{c2} = \frac{\mu_n C_{ox}}{2} \cdot (\frac{W}{L})_3 \cdot (V_{o1} - V_{tn3})^2 \tag{6.11}$$

$$I_{c1} = \frac{\mu_n C_{ox}}{2} \cdot (\frac{W}{L})_1 \cdot (V_{o2} - V_{tn1})^2 \tag{6.12}$$

$$I_{c2} = I_{c1} \tag{6.13}$$

Combining (6.10)–(6.13), V_{o1} and V_{o2} in circuit C in Fig. 6.1 can be solved as:

$$V_{o1} = \frac{[(1-\theta)\sqrt{\frac{(W/L)_3}{W_2/(L_2+L_6)}} - \sqrt{\frac{(W/L)_3}{(W/L)_1}}] \cdot V_{tn3} + (V_{tn1} - V_{tn2})}{(1-\theta)\sqrt{\frac{(W/L)_3}{W_2/(L_2+L_6)}} - \sqrt{\frac{(W/L)_3}{(W/L)_1}} + 1} \tag{6.14}$$

$$V_{o2} = \frac{V_{o1} - V_{tn3}}{\sqrt{\frac{(W/L)_1}{(W/L)_3}}} + V_{tn1} \tag{6.15}$$

It can be seen that the both circuits express CMOS threshold voltage. The combined analytical and numerical approach introduced in Chapter 5 can be applied to improve the design using circuits B and C topologies, and can achieve very high linearity with temperature at the output nodes.

IV. SIMULATION RESULTS AND PERFORMANCE SUMMARY

The simulation results in Fig.6.2 show the comparison among three circuit temperature non-linearities of node voltage V_{o2} at low V_{DD} (90% of nominal V_{DD}) and large threshold voltage corner. Circuits B and C are still able to maintain temperature error within 0.3°C while circuit A has more than 1°C temperature error.

Compared with node voltage V_{o1}, V_{o2} demonstrates wider voltage range than V_{o1} through the temperature range of -20°C to 100°C, and will present less burden on the

voltage range requirement of the circuit taking the sensor output as input. Therefore, in circuits B and C, the circuits are sized to obtain a very linear voltage with temperature at node V_{o2}. The overall performances of node V_{o2} in circuits B and C have been summarized in Table 6.1 and Table 6.2 respectively.

From Table 5.1, Table 6.1 and Table 6.2, circuit C has the best performance because the worst-case maximum temperature error is less than 0.11°C. This performance, to my knowledge, is the best performance in accuracy and robustness when compared with circuits A and B and other MOS based on-die temperature sensors [2] [3]. From the pattern of temperature errors over the corners between circuits B and C, it can be seen that high-order temperature non-linearity cancellation tracks better over the corners in circuit C than in circuit B.

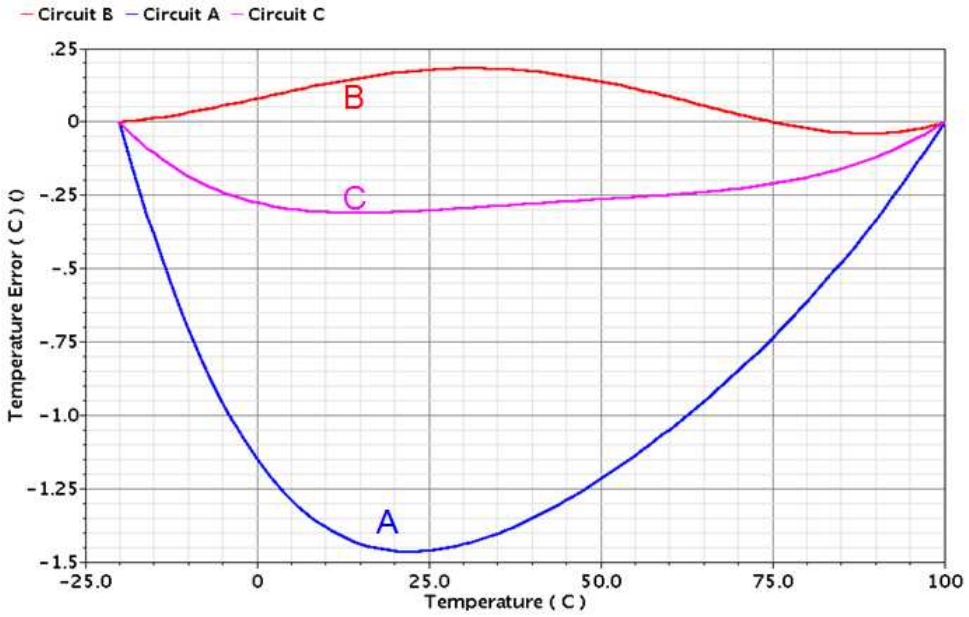

Figure 6.2 Temperature error curves of the three proposed circuits at the low V_{DD} high threshold corner

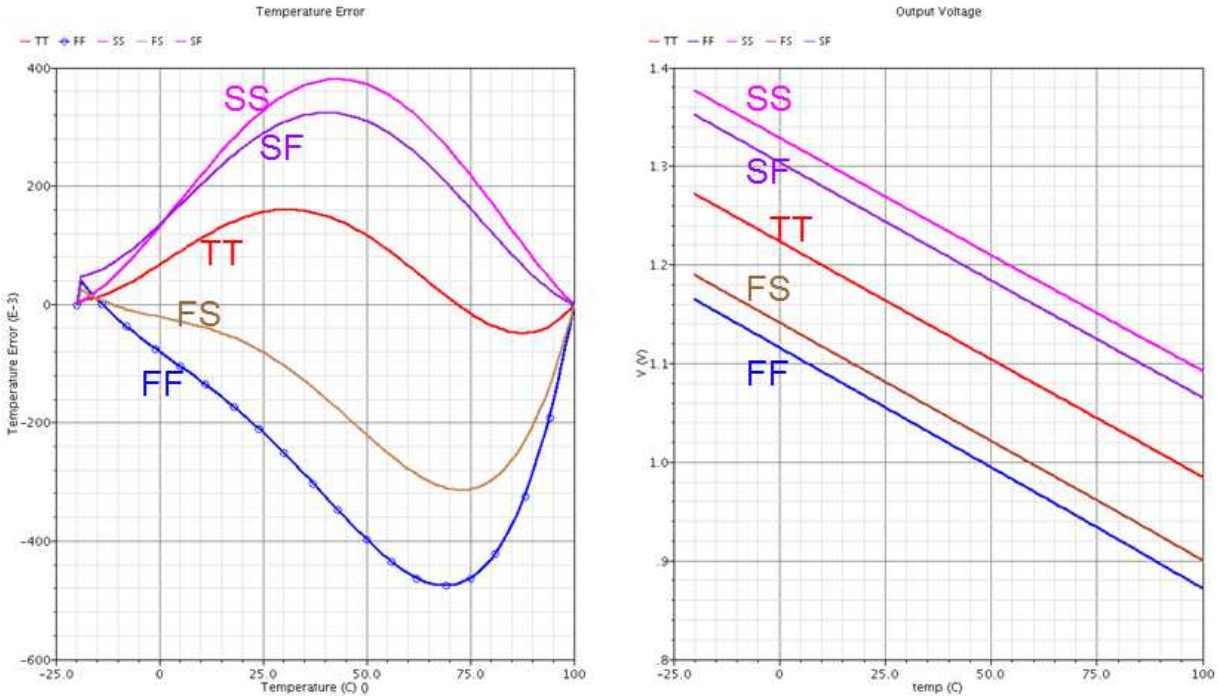

Figure 6.3 Circuit B temperature error curves and output voltage versus temperature transfer functions at various process corners

Table 6.1 Performance summary for circuit B

Parameters	*Performances*
Process (μm)	0.18
V_{DD} (V)	1.8
Temperature Range	-20 ~ 100
Maximum Temp Error (°C) at typical condition	0.14
Maximum Temp Error (°C) at worst corner	0.5
Maximum Temp Error (°C) at # 500 MC simulations	mean=0.17, std=0.05
Total gate area (μm^2)	270
Output Voltage Temp Coefficient (mV/ °C)	-2.39
Power consumption (μW)	95

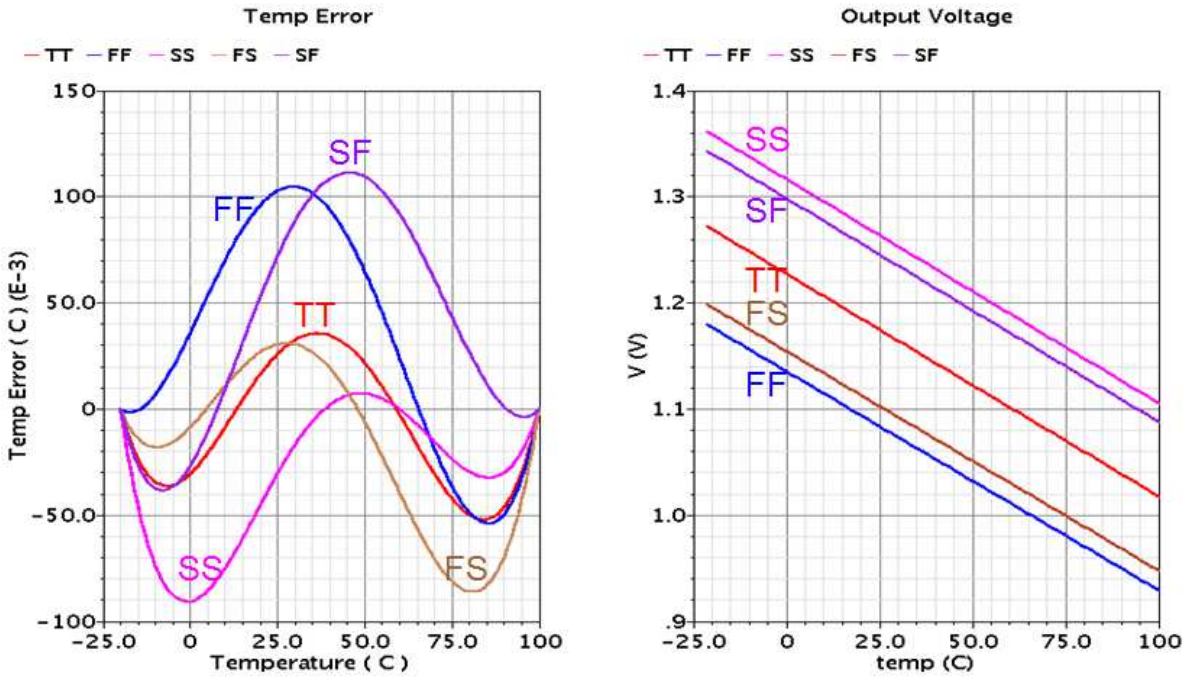

Figure 6.4 Circuit C temperature error curves and output voltage versus temperature transfer functions at various process corners

Table 6.2 Performance summary for Circuit C

Parameters	*Performances*
Process (μm)	0.18
V_{DD} (V)	1.8
Temperature Range	-20 ~ 100
Maximum Temp Error (°C) at typical condition	0.05
Maximum Temp Error (°C) at worst corner	0.11
Maximum Temp Error (°C) at # 500 MC simulations	mean=0.1, std=0.04
Total gate area (μm^2)	185
Output Voltage Temp Coefficient (mV/ °C)	-2.11
Power consumption (μW)	65

Table 6.3 Comparison with recent integrated temperature sensors

Sensors	Max Temp Error (°C)	Area (mm^2)	Temperature Range (°C)	Power Consumption	CMOS technology
Ring Oscillator based [2]	~7 (corners)	N/A (14 transistors)	40~150	25μW	45nm
Current-freq converter [3]	~ 1	0.003~0.02	0~120	200 μW	1μm
Ring oscillator based [6]	+2.75 ~ -2.9	0.0013	-40~110	400μW	65nm
TDC [7]	-0.7~+0.9	0.175	0~100	10μW	0.35μm
Thermal diode, Intel [4]	~±8	N/A	40~110	N/A	N/A
Thermal diode, AMD [11]	~±10	N/A	-55~85	10~100 μW	90nm
PN junction based[5] For PowerPC microprocessor	±4 with trim	N/A	10~100	N/A	N/A
PN junction based[8]	±1	~1.3	-40~120	7 μW	2 μm
PN junction based[9]	0.1 (one corner, 3σ)	4.5	-55~125	> 187.5 μW	0.7μm
CMOS threshold voltage based [12]	-1 ~ 0.8	50e-6 Needs External Current Mirror	50~125	25μW	90nm
Proposed (circuit B) (circuit C)	0.14(TT) 0.05 (TT)	~0.0013 ~0.0009	-20~100 -20~100	95μW 65μW	0.18μm

Table 6.3 lists the performances of published temperature sensors and compares the proposed temperature sensors. The proposed circuits in Chapter 6 have much better accuracy compared with other on-die temperature sensors for microprocessor thermal management applications [2]–[5]. Although the circuits do not have digital interface as [6]-[9], the advantages of the ultra small size and high linearity make the circuits very attractive as the analog sensing circuit part for temperature-to-digital converters.

V. CONCLUSIONS

In this chapter, the worst corner for circuit A introduced in Chapter 5 is identified. The reason for the temperature linearity degradation is analyzed. A simple approach of adding a triode region NMOS transistor is shown to be effective in reducing the

degradation at the worst corner—low V_{DD}, high threshold corner. Since this extra device is still a MOS transistor instead of a resistor, the output voltages are still extracting threshold voltage and suppressing mobility. After the modification, the circuits' performances show excellent linearity at typical conditions, as well as process corners, device mismatch and power supply variations.

REFERENCES

[1] J. Y. Kim and R. L. Geiger, "Characterisation of linear MOS active attenuator and amplifier." *Electronics Letter*, Vol. 31, pp. 511–513, 1995.

[2] B. Datta and W. Burleson, "Low-Power and Robust On-Chip Thermal Sensing Using Differential Ring Oscillators" *Proc. of IEEE Midwest Symposium on Circuits and Systems*, pp. 29-32, 2007.

[3] V. Szekely, C. Marta, Z. Kohari and M. Rencz, "CMOS Sensors for On-Line Thermal Monitoring of VLSI Circuits", Vol.5, *IEEE Trans. On VLSI Systems*, pp. 270-276, September 1997.

[4] Intel Corporation, "*Mobile Intel 965 Express Chipset Family Datasheet, Revision 3*", June 2007.

[5] H. Sanchezm, et al., "Thermal Management Systems for High Performance Power PC Microprocessors", *Proc. IEEE Compcon*, pp. 325-330, Feb. 1997.

[6] K. Kim, H. Lee, S. Jung, C. Kim, "366kS/s 400uW 0.0013mm2 Frequency-to-Digital Converter Based CMOS Temperature Sensor Utilizing Multiphase Clock," *IEEE Custom Intergrated Circuits Conference*, pp. 203-206, 2009.

[7] P. Chen, C. Chen, C. Tsai, and W. Lu, "A Time-to-Digital-Based CMOS Smart Temperature Sensor", *IEEE Journal of Solid State Circuits*, Vo1. 40, pp. 1642-1648, August 2005.

[8] A. Bakker and J. Huijsing, "Micropower CMOS Temperature Sensor with Digital Output", *IEEE Journal of Solid State Circuits*, Vol. 2, pp. 933-937, July 1996.

[9] M. Pertijs, K. Makinwa, and J. Huijsing, "A CMOS Smart Temperature Sensor with a 3σ Inaccuracy of ±1^{o}C from -55^{o}C to 125^{o}C", *IEEE Journal of Solid State Sircuits*, Vol.40, pp. 2805-2815, Dec. 2005.

[10] M. Sasaki, M. Ikeda, M. Asada, "A temperature sensor with an in- accuracy of -1/+0.8◦C using 90-nm 1-V CMOS for online thermal monitoring of VLSI circuits", IEEE Trans. on Semiconductor Manufacturing, Vo1. 21, pp.201-207, May 2008.

[11] Advanced Micro Devices Inc, "*AMD Functional Data Sheet, 940 Pin Package, Rev 3.05*", June 2004

[12] M. Sasaki, M. Ikeda, M. Asada, "A temperature sensor with an in- accuracy of -1/+0.8◦C using 90-nm 1-V CMOS for online thermal monitoring of VLSI circuits", IEEE Trans. on Semiconductor Manufacturing, Vo1. 21, pp.201-207, May 2008.

CHAPTER 7
MASTER-SLAVE HYBRID ON-DIE TEMPERATURE TO DIGITAL CONVERTER

ABSTRACT

In this chapter, a new method is introduced for on-die temperature measurements that directly produces digital output codes that are highly linear in temperature. Unlike conventional bipolar junction tansistors (BJTs) based temperature sensor circuits with digital readouts, all of which employ distinctive temperature sensing elements, signal conditioning and amplification circuits, accurate temperature-stable reference circuits, and a sufficiently linear analog to digital converter, the proposed architecture steers a current into two temperature sensitive nodes through digital-to-analog converter (DAC) control. When the two node voltages become equal, the DAC code gives a digital readout of the on-die temperature. This new structure does not require a reference circuit, and the complexity of implementation is relaxed, yet it still provides very good temperature linearity. Unlike the inverter delay based structures that are sensitive to process variations and require extra area-consuming support hardware to compensate for the process variations, the proposed topology can operate at nominal as well as four other standard process corners and still maintain good temperature linearity. In this new topology, the DAC current is generated from a simple V_{DD} independent current generator and does not need to be constant or linear with temperature. The DAC control code is produced by a successive approximation register (SAR) logic circuit that is driven by a comparator that compares the two node voltages. When the two outputs differ, the SAR is repeatedly updated to decrease the difference between the two node voltages. When the two node

voltages are essentially the same, updating of the SAR is inhibited and the boolean output is the digital code that resides in the SAR. A sizing strategy based upon a combined analytical and numerical optimization approach is used to provide excellent linearity of the digital code with respect to temperature. Simulaton results demonstrate that the maximum nonlinear temperature error can be held, with modest design efforts, to less than 0.5^{o}C over a 120^{o}C temperature range over all process corners and a ±10% power supply variation.

I. INTRODUCTION

Research on integrated temperature sensors started in the mid-seventies [1]. One of the drivers for the temperature sensor development was the need for convenient digital interfaces to provide ready access to on-die temperature by digital circuits. As Systems on Chip (SoC) become mainstream and more traditional analog blocks (such as phase locked loops (PLLs)) are replaced with digital counterparts, the need for digital representations of die temperature becomes more pressing. Thanks to the relentless drive of the semiconductor industry toward very large scale integration in the past four decades, analog-to-digital conversion and bus interface circuits are becoming smaller and smaller so that they can be integrated with the temperature sensor. In 1992, Smartec produced the first temperature sensor with a duty-cycle output [2]. In 1995, National Semiconductor unveiled its temperature sensor chip, LM75, with a fully integrated bandgap temperature sensor core, a sigma-delta analog-to-digital converter and an I^2C interface [3].

A major milestone in temperature sensors development came in the late nineties when Intel added temperature sensors to their microprocessors and motherboards. This

marked the beginning of the age when chip and/or die temperature measurements with digital output representations of the temperature are an integral part of the design of large integrated circuits and systems that rely on accurate measurements of the die temperatures to maintain the desired reliability and performance of the systems. These real-time, on-die temperature measurements provide critical sensory data to the power/thermal mangement circuits that are necessary to achieve accpetable reliability and performance optimization. This approach makes it possible for the performance of a microprocessor to be improved at the architectual and the operating systems levels, as an alternative to simply increasing clock frequencies.

Although accurate measurement of temperature using on-chip temperature monitors appears to be a straightforward task, the accuracy of most integrated temperature monitors is not sufficient. The area of existing temperature sensors with good accuracy is often large, and considerable processing overheads are often required to convert the temperature measurements to the digital form required by the power management controller (PMC) in a large digital system.

Considering the applications of temperature measurements in the design of microprocessor systems, the design objectives are not only to have good accuracy, a small area, and low power design, but also to provide a digital output for the convenience of communicating with other VLSI blocks in the complex system.

Considerable research has been associated with on-chip temperature sensors with digtal outputs. Generally speaking, there are two classes of temperature sensors with digital output: analog BJTs based structures [4]-[6], and inverter delay line based structures [8]-[11] [18]. In an analog BJTs based structure, a temperature sensor core is

used to generate a voltage or current that is linearly dependent on temperature, and either a delta-sigma analog-to-digital converter (ADC) or a SAR ADC is then used to convert the analog voltage or current into a digital code. This approach usually requires large area and high power consumption. Fig.7.1 shows a block diagram of the standard approach.

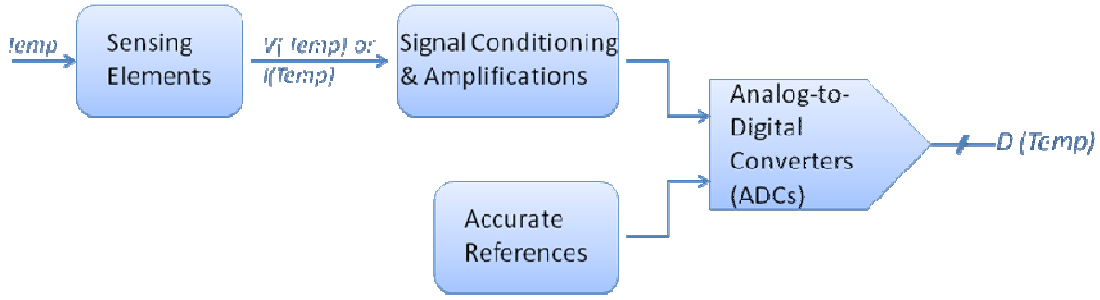

Figure 7.1 Conventional temperature to digital sensor topology

This conventional temperature-to-digital converter has four distinct blocks:

(1) A temperature sensing element that generates a voltage or a current that linearly changes with temperature;

(2) A signal conditioning and amplification block that adjusts the voltage or current quantity to an acceptable level with sufficient swing to be compatible with the input signal range of the next stage ADC;

(3) An accurate and thermally stable reference, which is usually built by pn junctions and an op amp, providing a reference voltage or reference current for the ADC;

(4) An linear ADC, usually a sigma-delta modulator based ADC (ΣΔADC) or a successive approximation register based ADC (SAR ADC).

In this standard approach, to guarantee a certain level of linearity in the relationship between temperature and the digitized output, conventional wisdom suggests that each of the four blocks must have linearity or accuracy beyond what is required of

the whole system: the temperature sensing element must have an output linearly dependent upon temperature; the signal shifting and amplifying circuit must be linear, and the gain and offset must be independent of temperature; the reference must be accurate and temperature stable, and the ADC must be linear and stationary with temperature. If any of these requirements are violated, the digital sensor output cannot be guaranteed to be accurate. With this approach, to reduce the chip-level design challenges, some designers rely on the an external accurate reference to achieve the desird accuracy level [6]. Other use complicated techniques such as curvature correction to improve the thermal stability of the bandgap voltage references [7]. But, from a practical viewpoint, there is considerable incentive to have the temperature to digital converter completely integrated.

The conventional temperature sensor topology in Fig.7.1 can be viewed as an open loop system. To preserve the linear relationship between temperature and the final digital output, the designs of each block in this system are quite challenging if the required accuracy is at the 0.5^{o}C level. Reported temperature sensors using such a topology are typically able to achieve 1^{o}C accuracy or better with a good analog design and with calibration. By combining sophisticated analog techniques, such as offset cancellation and dynamic element matching with oven-controlled accurate calibration, an accuracy of $\pm 0.1^{o}$C was reported [8] recently but the reported area required to attain thie level of performance was 4.5 mm^2 and is orders of magnitude too large for embedded applications. By contrast, the new topology we propose in this chapter is a closed-loop feedback system and does not need any thermally stable reference, does not require any

linear ADC, and does not require any accurate signal amplification or signal conditioning circuits.

The second class of structures that have been reported is based upon the temperature-dependent inverter delays of boolean inverters. There are three reported methods for sensing temperature using temperature-dependent inverter delays: a loop delay line, a delay locked loop (DLL), and a ring oscillator. In the loop delay line method, a cyclic time-to-digital converter converts a pulse with a width proportional to the temperature into a corresponding digital code [9]–[12]. To achieve sufficient temperature resolution, hundreds of inverters are required to obtain enough pulse delay. Unfortunately, the inverter delays are adversely affected by process variations of threshold voltage V_{th}, mobility μ and gate oxide capacitance C_{ox}. The DLL method presented in [19] uses a delay locked loop (DLL) nested in another DLL to measure temperature. Each delay line needs a multiplexer and a phase interpolator for selecting the desired amount of delay. The reported DLL-based temperature sensor occupies a large area around 0.16mm^2, and consumes high power at about 1.2mW. The third method of inverter delay based temperature sensors uses frequency changes in current-starved ring oscillators as an indicator of temperature. With this approach, one ring oscillator serves as a reference and another ring oscillator has an oscillation frequency dependent on temperature. The main advantages of the ring oscillator method are the very small area and the scalibility with technology. One limitation of the current-starved ring oscillator approach is the need for accurate temperature-independent biasing currents [20]. As with any inverter delay based approach, the ring oscillator output is affected by both mobility and threshold voltage, resulting in significant temperature nonlinearity.

In this chapter, we propose a new approach for on-chip temperature measurement that provides a direct digital readout. A specific implementation of this method is depicted in Fig.7.2. It is comprised of a simple V_{DD} independent current generator circuit called the "master circuit" and a feedback loop comprised of a slave temperature sensitive sensor circuit blended with a current-steering DAC, a comparator and a SAR logic block. It does not require any conventional constant reference voltage or reference current, and it does not attempt to make any node voltage or branch current constant or precisely linear to temperature. It does not need any conventional ADCs, yet it generates a digital output code that is extremely linear to temperature. This new topology, together with our sizing strategy that combines analytical and numerical optimization similar to that discussed in Chapter 5, enables us to implement a temperature to digital converter with linearity of about 0.2^{o}C over a wide temperature range yet area required for the implementation is small making it suitable for on-die multi-site measurement applications.

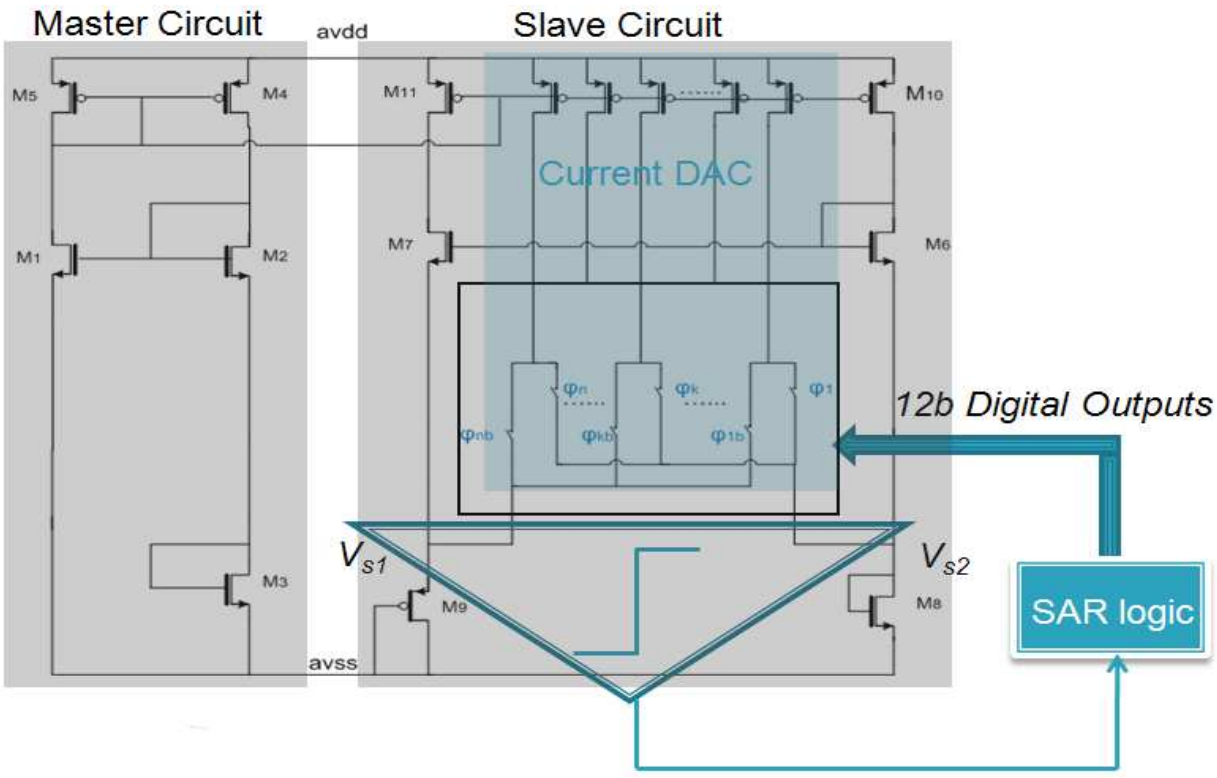

Figure 7.2 Diagram for the proposed master-slave hybrid on-die temperature to digital unit

II. TOPOLOGY OF MASTER SLAVE HYBRID ON-DIE TEMPERATURE SENSOR

Transistor-level details of one method for implementing the new approach to building on-die temperature sensors, shown in Fig. 7.2, will be discussed in this section. The proposed temperature measurement structure is comprised of three major parts. One is termed the"master circuit", which can be viewed a traditional V_{SS} independent voltage or current generator. It consists of the five transistors M_1–M_5. It provides a bias voltage for the second block—"slave circuit." If a basic square-law model is used for the devices and output conductance and bulk modulation effects are neglected, the master circuit "expresses" the threshold voltage. The master circuit does not need to provide a constant voltage and, to the contrary, the temperature dependence of the output voltage it

generates is an integral part of the temperature sensor. The second block consists of a simple current steering DAC comprised of transistors M_{d1}……M_{dn}, and a slave temperature sensitive circuit made by transistors M_6-M_{11}. The SAR logic and a comparator create a feedback path and form the third block. The whole circuit operates in the following way: at each temperature, there is an initial voltage difference between the sensing nodes V_{S1}, and V_{S2}. The comparator will compare this voltage difference, and based upon the boolean state of the comparator, drive the SAR logic to find a digital code θ that forces the difference between V_{S1} and V_{S2} to be very small. The difference between these two voltages is affected by the current of the DAC which has the SAR output as an address. The DAC steers a portion of the total current into M_9 and the remaining current into M_8, so that at the end of the conversion cycle, condition (7.1) will be approximately satisfied. How closly (7.1) is satisfied is dependent upon the offset voltage of the comparator and the resolution of the current-steering DAC:

$$V_{s1} = V_{s2} \tag{7.1}$$

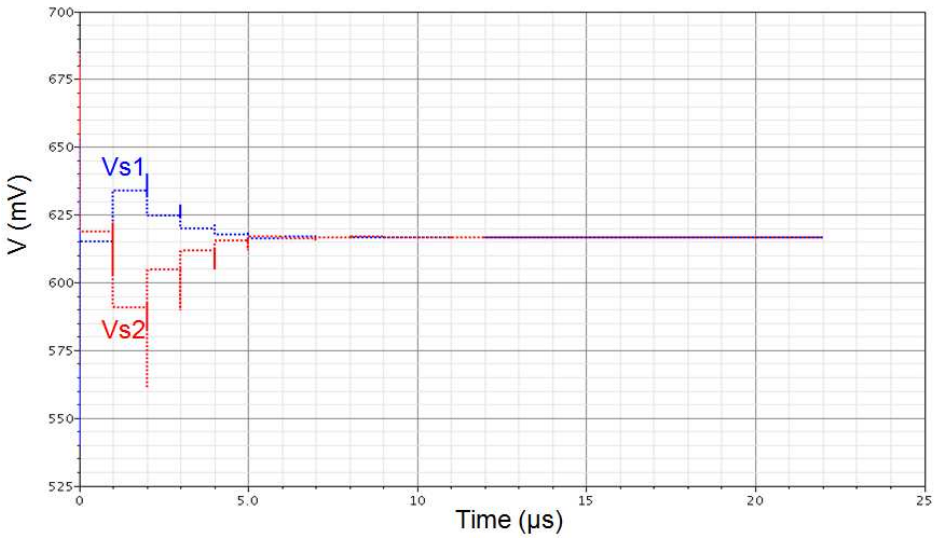

Figure 7.3 Transient voltages at the two sensing nodes during a conversion cycle

Fig. 7.3 shows results from a computer simulation of the voltages V_{S1} and V_{S2} as they converge when a binary search is uesd to control the SAR. The temperature dependence of the digital code can be estimated from the following analysis. Assume the current in each branch of the master circuit is I_{ref}, and the current mirror ratio between M_4, M_{11} and M_{10} is 1: m: m.

Therefore,

$$I_{M9} = I_{M8} \tag{7.2}$$

The current DAC is binarily weighted and has a total current of $n \bullet I_{ref}$. At the end of the conversion cycle, by applying KCL (Kirchoff's Current Law) at the two sensor node V_{S1} and V_{S2}, we can obtain the following equations:

$$I_{M9} = m \cdot I_{ref} + \theta \cdot n \cdot I_{ref} = \frac{1}{2}\mu_p C_{ox}(\frac{W}{L})_9(V_{S1} - |V_{tp9}|)^2 \tag{7.3}$$

$$I_{M8} = m \cdot I_{ref} + n \cdot I_{ref} \cdot (1-\theta) = \frac{1}{2}\mu_n C_{ox}(\frac{W}{L})_8(V_{S2} - |V_{tn8}|)^2 \tag{7.4}$$

where θ is the analog interpretation of the digital code normalized to a full scale value of one; I_{ref} is generated by the "master circuit" and, using the square law model of the devices, can be expressed as:

$$I_{ref} = \frac{1}{2}\mu_n C_{ox}(\frac{W}{L})_3(V_{o1} - V_{tn3})^2 \tag{7.5}$$

where V_{o1} is the node voltage at the drain of M_3 in the master circuit. This voltage was derived explicitly and was given in (5.7) of Chapter 5.

Combining (7.1)-(7.5), it follows that when (7.1) is approximately satisfied, (7.6) is an implicit equation involving θ. This equation can be solved to obtain an expression for θ.

$$\sqrt{\frac{(m+n\cdot\theta)\cdot(W/L)_3}{(W/L)_9}\frac{\mu_n}{\mu_p}}-\sqrt{\frac{(m+n\cdot(1-\theta))\cdot(W/L)_3}{(W/L)_8}}=\frac{V_{tn8}-|V_{tp9}|}{V_{01}-V_{tn3}} \tag{7.6}$$

Solving θ from (7.6) gives the expression:

$$\theta=\frac{(n+m)\cdot V_{tn}^2-ma^2\cdot V_{tn}^2-b^2\cdot\Delta V_t^2}{V_{tn}^2(a^2+1)n} +\frac{2\cdot a\cdot b\cdot\Delta V_t\cdot(ab\cdot\Delta V_t+\sqrt{2m(a^2+1)\cdot V_{tn}^2+na^2\cdot V_{tn}^2+n\cdot V_{tn}^2-b^2\cdot\Delta V_t^2})}{V_{tn}^2\cdot(a^2+1)^2 n} \tag{7.7}$$

where

$$a=\sqrt{\frac{\mu_n}{\mu_p}\cdot\frac{(W/L)_8}{(W/L)_9}} \tag{7.8}$$

$$b=\sqrt{\frac{(W/L)_8}{(W/L)_3}}\cdot\frac{\sqrt{\frac{(W/L)_2}{(W/L)_1}}-\sqrt{\frac{(W/L)_2}{(W/L)_3}}-1}{\sqrt{\frac{(W/L)_2}{(W/L)_3}}} \tag{7.9}$$

In (7.8), the mobility μ has a temperature dependence modeled by the following expression [21]:

$$\mu=\mu_0(T_0)\cdot(T/T_0)^{ute} \tag{7.10}$$

where T_0 is the reference temperature, and *ute* is a negative constant dependent upon the process.

In (7.7), ΔV_t is the difference of threshold voltage of NMOS transistor M_8 and PMOS transistor M_9. All the NMOS transistors are assumed to have the same threshold voltage V_{th}. The temperature dependence of the threshold voltage of an nMOS transistor was given in equation (5.9) of Chapter 5 which is repeated as

$$V_{tn}(T) = V_{tn0} + (KT1 + KT1L / L_{eff} + KT2 \cdot V_{bseff}) \cdot (T / T_{NOM} - 1) \qquad (7.11)$$

If the bulk voltage effect is neglected and if it is assumed that L_{eff} is not temperature dependent, it follows that $V_{tn}(T)$ can be expressed as

$$V_{tn}(T) = \left(V_{tn0} - (KT1 + KT1L / L_{eff})\right) + \left(\frac{KT1 + KT1L / L_{eff}}{T_{NOM}}\right) T \qquad (7.12)$$

where the terms in parenthesis on the right-hand side of this equation are independent of temperature. The linear dependence of the threshold voltage with temperature should be apparent. For derivation convenience, this can be expressed as

$$V_{tn}(T) = V_{tn00} - \alpha_n T \qquad (7.13)$$

where V_{tn00} and α_n can be viewed as process parameters that are independent of temperature and given by the expressions

$$V_{tn00} = \left(V_{tn0} - (KT1 + KT1L / L_{eff})\right) \qquad (7.14)$$

and

$$\alpha_n = \left(-\frac{KT1 + KT1L / L_{eff}}{T_{NOM}}\right) \qquad (7.15)$$

The term V_{tn00} is the 0K intercept of the temperature dependence of the threshold voltage and α_n is the temperature coefficient of the threshold voltage. Both V_{tn00} and α_n are positive for n-channel transistors.

Correspondingly, the threshold voltage for p-channel transsistors can be expressed as

$$V_{tp}(T) = V_{tp00} + \alpha_p T \tag{7.16}$$

where V_{tp00} and α_P can be viewed as process paramaters that are independent of temperature. The term V_{tp00} is the 0K intercept of the temperature dependence of the threshold voltage. V_{tp00} is negative and α_p is positive for p-channel transistors.

Since ΔV_t is much smaller than V_{tn}, and the circuit, which will be shown in Section III, is sized in such a way that in (7.7) the ΔV_t^2 terms can be neglected. Then,

$$\theta = \frac{(n+m) - ma^2}{(a^2+1)n} + \frac{\Delta V_t}{V_{tn}} \cdot \frac{2 \cdot a \cdot b \cdot \sqrt{2m(a^2+1) + na^2 + n}}{(a^2+1)^2 n} \tag{7.17}$$

Combining (7.13), (7.16) and (7.17), the following expression can be obtained:

$$\theta = \left[\frac{(n+m) - ma^2}{(a^2+1)n} \right] + \frac{(V_{tn00} - |V_{tp00}|) - (\alpha_n - |\alpha_p|)T}{V_{tn00} - \alpha_n T} \cdot \left[\frac{2 \cdot a \cdot b \cdot \sqrt{2m(a^2+1) + na^2 + n}}{(a^2+1)^2 n} \right] \tag{7.18}$$

It follows from (7.8) and (7.10) that the term *a* is only weakly dependent upon temperature. If this temperature dependence is neglected, the two terms in brackets in (7.18) are independent of temperature. If the term V_{tn00} is large compared to $\alpha_n T$, the denominator in the second term on the right hand side of (7.18) is also only weakly dependent on temperature and the paramater θ becomes nearly lineary dependent on temperature. This simplified analysis shows that the dependence of θ with temperature is reasonably linear and further investigation is required to determine how much temperature nonlinearity is introduced by using the simiplified model in the analysis and by neglecting the weakly temperature-nonlinear terms in (7.7). This willl be addressed

later in this section using computer simulation results that are obtained by using more detailed device models.

From (7.18) it can be observed that when normal process varations are present, the ΔV_{th} term could change significantly thus causing rather large swings in θ. And, if the magnitude of the nominal n-channel and p-channel threshold voltages are comparable, there could even be a sign change in θ. Such a large variation in θ due to process variations is not attractive. Also, if ΔV_{th} were to vanish, the paramater θ would loose its temperature dependence. Neither scenario is attractive. Since the threshold voltage is a function of the device length, it would be advisable to choose the lengths of M_8 and M_9 so that the change in θ with process variations is not too large and so that it does not change signs.

From (7.8), (7.9) and (7.18), it can be observed that the temperature dependence of the normalized digital code is directly related to the aspect ratio of transistors M_1–M_3, M_8–M_9 and that there is some nonlinearity in the relationship between temperature and θ. The question about whether the sizing of M_1–M_3, M_8, M_9 can be used to improve linearity naturally arises. The analysis needed to obtain (7.18) was obtained under the assumption that output conductance effects (λ effects) are negligable. It can be shown that λ effects in the circuit affect the analysis and introduce some additional temperature nonlinearities. Inclusion of λ effects in an analytical formulation needed to obtain a more exact expression for θ would be very cumbersome and no attempt will be made to develop such a forumlation in this work.

Regardless of the exact formulation, the boolean output represented by the paramater, θ still changes with temperature. The resulting nonlinearity in θ with respect

to temperature can be compensated by using a numerical approach similar to that to that demonstrated in Chaper 5 to adjust devices sizes M_1–M_3, M_8, M_9. The performances of the temperature sensor based upon the numerical optimization are discussed in the following section.

III. DESIGN EXAMPLE AND SIMULATION RESULTS

An implementation of the temperature sensor shown in Fig.7.2 is discussed in this section. In this implementation, focus has been placed on the analog front end portion of the circuit, including the master circuit, the slave sensor and the PMOS current DACs. These blocks form the sensor core and have the most direct influence on the output digital code temperature linearity. A 12-bit digital code from the SAR logic is used to steer current in the PMOS current mirror. The 12-bit resolution is selected so that there is enough current steering capability to resolve the temperature difference to the 0.05^{o}C levels over all process corners over the temperature range -20^{o}C to 100^{o}C. At any particular corner, the typical corner for example, 10-bit of resolution is needed for the DAC. The additional two bits of resolution are needed to provide operation over the standard process corners. The digital SAR logic is implemented in VerilogA since the circuit-level details of the SAR logic should have negligable impact on the temperature non-linearity of the sensor. Comparator offset can potentially cause errors in the digital output code. However, there are standard approaches, such as digital calibration or the use of an autozeroing technique to keep the offset around or under the 200μV level [16]–[18]. The tolerance of the comparator offset can be estimated as follows: Assume the LSB current of 12-bit current DAC applied is around 20nA and the diode connected

devices M_8 and M_9 have an output impedence of around 5KΩ. To be able to detect a temperature difference at the 0.2^oC level, the offset from the comparator needs to be below 200μV. The offset constraint can be relaxed by increasing the total current consumption in the PMOS current mirror or by increasing the output impedence of the diode connected transistors M_8 and M_9.

In the analog front end, two main sources can cause a temperature non-linearity of the digital output: the mismatch between outputs of the current mirror; and the temperature non-linearity from the master circuit and the slave sensors. The first error source can be reduced by scaling up the size of the PMOS current mirror. The second error, according to design experience, can be reduced by adjusting the size of devices M_1–M_3, M_6, M_8, M_9 and the total current consumption in the PMOS current mirrors relative to the bias current in slave circuit. The analog front is implemented in 0.18μm process technology with a 1.8V single ended power supply. An optimization following the basic approach discussed in Chapter 5 was used to size all devices. The sizing of the devices in the analog portion of the sensor are summarized in Table 7.1. In this optimization, a standard BSIM 3v3 model provided by the vendor for the process was used. This same BSIM 3v3 model was used in all simulations of this circuit. The circuit simulator program, SPECTRE, included as a part of the CADENCE toolset that is available on the Iowa State Univrsity campus, was used for all circuit-level simulations.

Table 7.1 Analog front-end dimensions

Master Circuit	Slave Sensor	PMOS current mirror (Binary weighted)
$(W/L)_1$=0.3 µ / 0.8 µ $(W/L)_2$=2.2 µ / 0.4 µ $(W/L)_3$=2x2 µ / 0.3 µ (for TT, SS, FF corners) =1x2 µ / 0.3 µ (for SF, FS) $(W/L)_4$=64x4.5 µ / 0.9 µ $(W/L)_5$=64x4.5 µ / 0.9 µ	$(W/L)_6$=2x5 µ / 0.5 µ $(W/L)_7$=5 µ / 2 µ $(W/L)_8$=8 µ / 0.6 µ $(W/L)_9$=40 µ / 1 µ $(W/L)_{10}$=60x3 µ / 0.9 µ $(W/L)_{11}$=60x3 µ / 0.9 µ	$(W/L)_{MSB}$=32x4.5 µ / 0.9 µ $(W/L)_{LSB}$=0.22 µ / 3.6 µ $(W/L)_{switch\_MSB}$=32x1µ / 0.22 µ $(W/L)_{switch\_LSB}$=0.25µ / 0.22 µ

The simulated output digital codes versus temperature, normalized to a reference of 1V, over the -20°C to 100°C temperature range for the design, summarized in Table 7.1, are shown in Fig. 7.4 assuming typical process paramaters. For reasons discussed above, instead of utilizing the full range of the 12-bit code, the digital output utilizes only about 20% of the full scale range at typical condition, covering from binary code range 010100110111 to 100001011010. From Fig. 7.4, it can be seen that the Boolean output is very linear with respect to temperature.

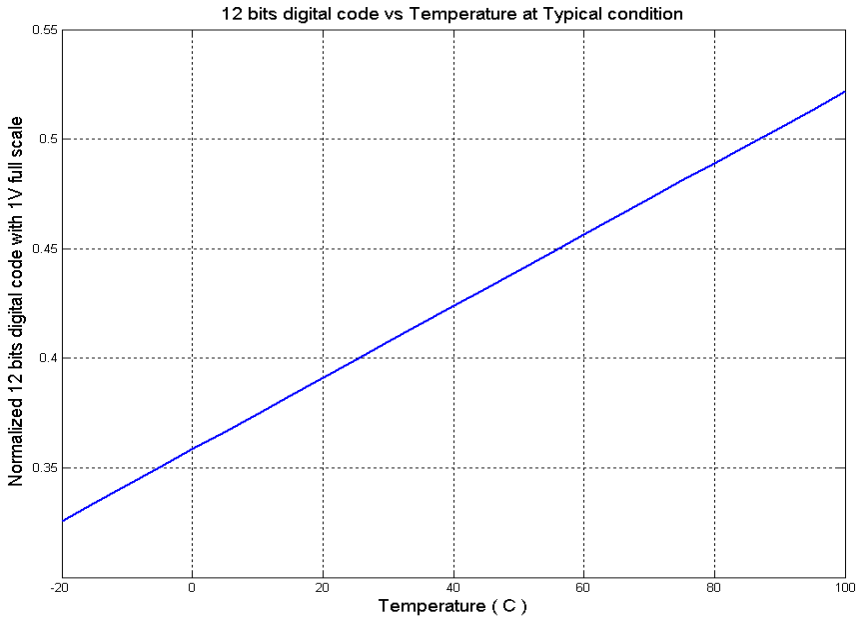

Figure 7.4 Normalized digital output code vs. temperature

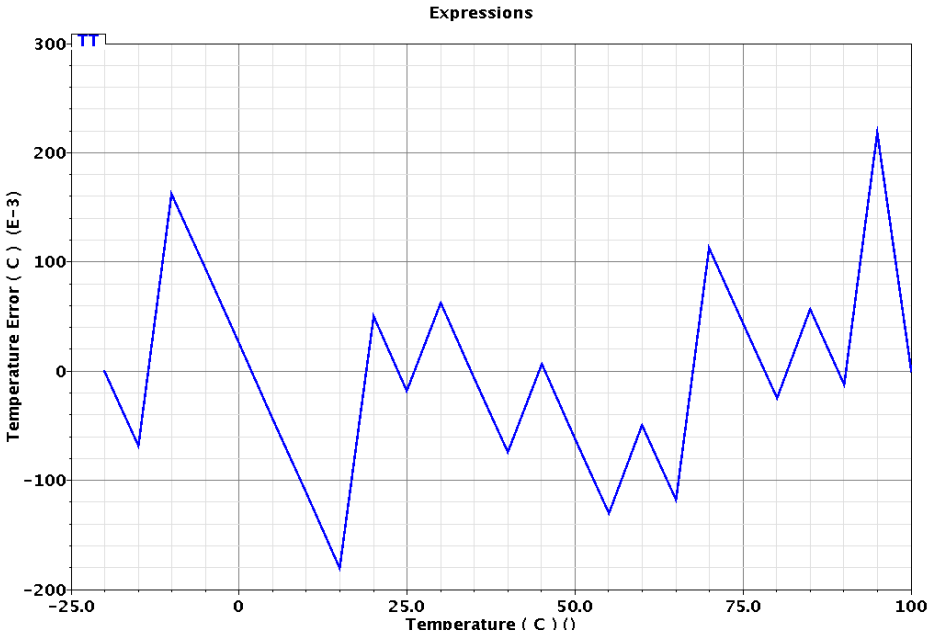

Figure 7.5 Temperature error of the digital codes at typical conditions

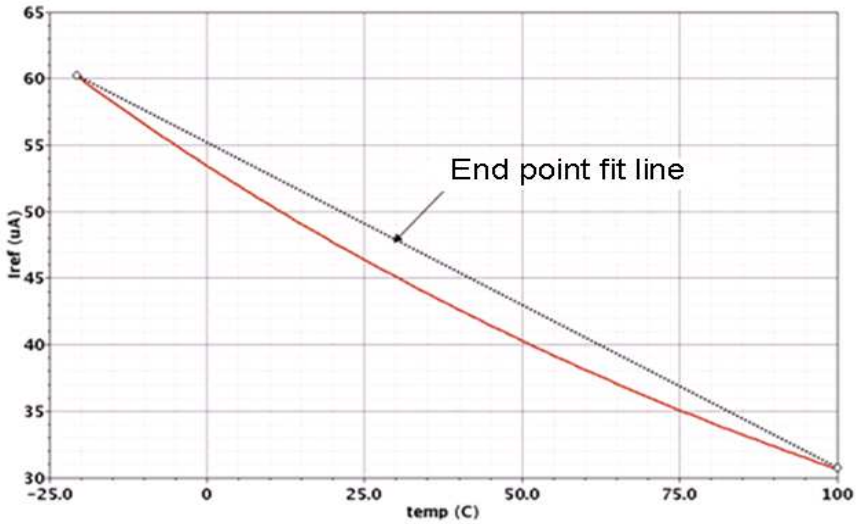

Figure 7.6 Reference current thermal stability and temperature nonlinearity

The digital outputs at all five process corners, normalized to a 1V reference, are shown in Fig.7.8. This shows why the extra two bits were needed to allow operation over all process corners.

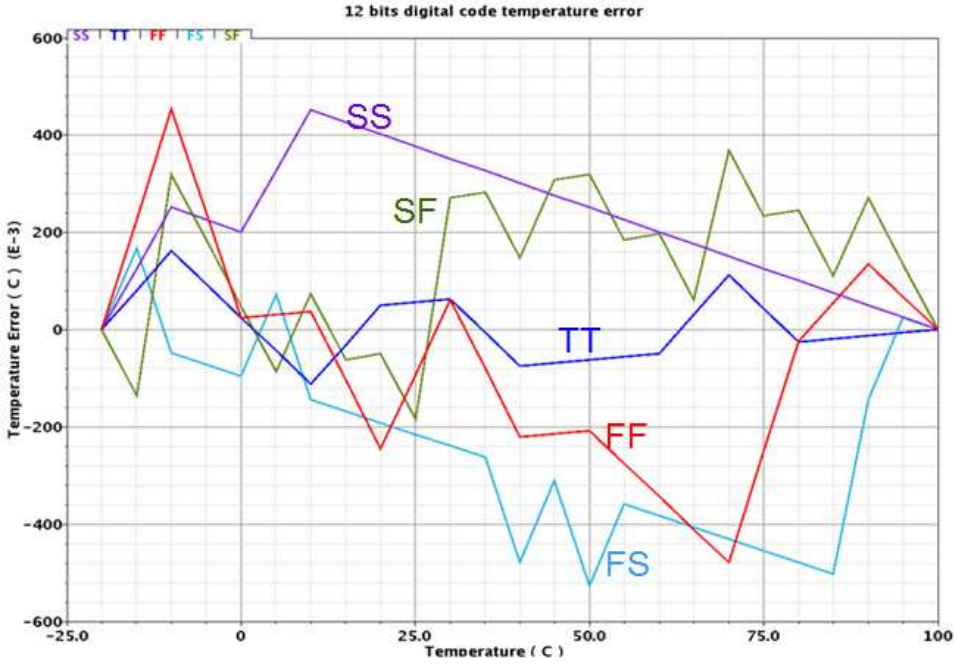

Figure 7.7 Temperature error of 12-bit digital codes over process corners

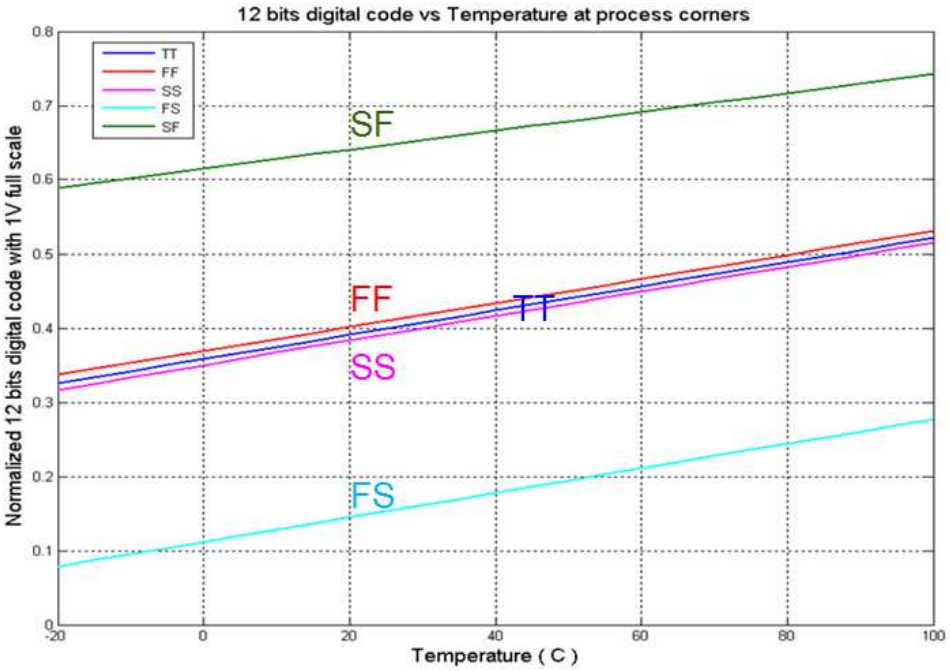

Figure 7.8 12-bit digital codes vs. temperature over process corners

Because of the highly linear output, the nonlinearity of the temperature characteristics cannot be seen in Fig. 7.4. In Fig. 7.5, the nonlinearity in the temperature sensor, still under the assumption of operating at the typical corner of the process, is shown. With the 12-bit output, simulations show a 0.2^{o}C temperature error over the whole temperature range. The temperature error was obtained by first determining the output code from -20^{o}C to 100^{o}C with 5^{o}C step and then by comparing the transfer curve against the end point fit line. The plotting program connected points in Fig. 7.5 but no significance should be placed at any temperatures increment other than at the 5^{o}C increments used in the simulation.

Simulation results for the same circuit over process corners are shown in Fig. 7.8. The term "same circuit" is used to indicate that the device sizes given in Table 7.1 remain the same without re-optimization for operating at different points in the process. A deterministic 1-bit calibration was undertaken to accommodate for operating at extremes in the process. In this figure, five different process points (corners) are included. In addition to typical, process corners fastNfastP, fastNslowP, slowNfastP and slowNslowP were used. From these simulations, it can be seen that output code remains quite linear over temperature at different corners of the process though there is a shift in the digital output code range. At the extreme process corners used in this simulation, a deterministic 1-bit calibration (tuning) element was added in paralled with the device M_3 in the master circuit. For typical, fastNfastP, slowNslowP corners, M_3 with an aspect ratio of 2x2μm/0.3μm was used while at the fastNslowP, slowNfastP corners, M_3 with a halved aspect ratio was used. The hardware overhead associated with this 1-bit tuning is very small.

The nonlinearity relative to end-point fit lines at all five process corners is shown in Fig. 7.7. Again, in this figure, simulation results were made with 5^{o}C increments, and also no signifcance should be placed on interpreted values associated with connecting the simulation points in the figure. From this figure, it can be observed that the nonlinear error is below 0.5^{o}C over the all five process corners.

The reference current provided by the master circuit is temperature dependent and contains visible temperature non-linearity as predicted by (7.5). Simulated results for the relationship between the reference current and tempearture at the typical process point are shown in Fig. 7.6. From this figure, it is apparent that the relationship between bias current and temperature is highly nonlinear but this nonlinearity is not reflected in a nonlinearity in the digital output code. This figure shows that the proposed structure does not require a conventional current reference with good thermal stability.

Most data converters require a reference, whether it is a voltage reference or a current reference. The proposed circuit of Fig. 7.2 shows no voltage reference and no current reference. And, although the term DAC was used to describe the switchable p-channel part of the slave circuit, the circuit is more a current steering array controlled by a SAR than it is a DAC. Indeed, this DAC has no reference and neither an output current nor an output voltage. This SAR-based current steering structure is rather embedded in a thermal self-compensating feedback loop that forces two node voltages to be equal.

Although no reference voltge is required, the circuit does require a supply voltage. The master circuit can be viewed as a V_{DD} independent current generator and the slave circuit is ideally V_{DD}-independent as well. But, since the output conductances of the devices are not 0, it is expected that V_{DD} will have some effect on the output. The

nonlinear error for the temperature sensor at the nominal supply voltage of 1.8V is compared with that obtained with ±10% power supply variations in Fig. 7.9. From this figure, it can be observed that the temperature error is less than 0.3°C with ±10% power supply variations. This suggests that the temperature-to-digital converter also shows good robustness to supply variations.

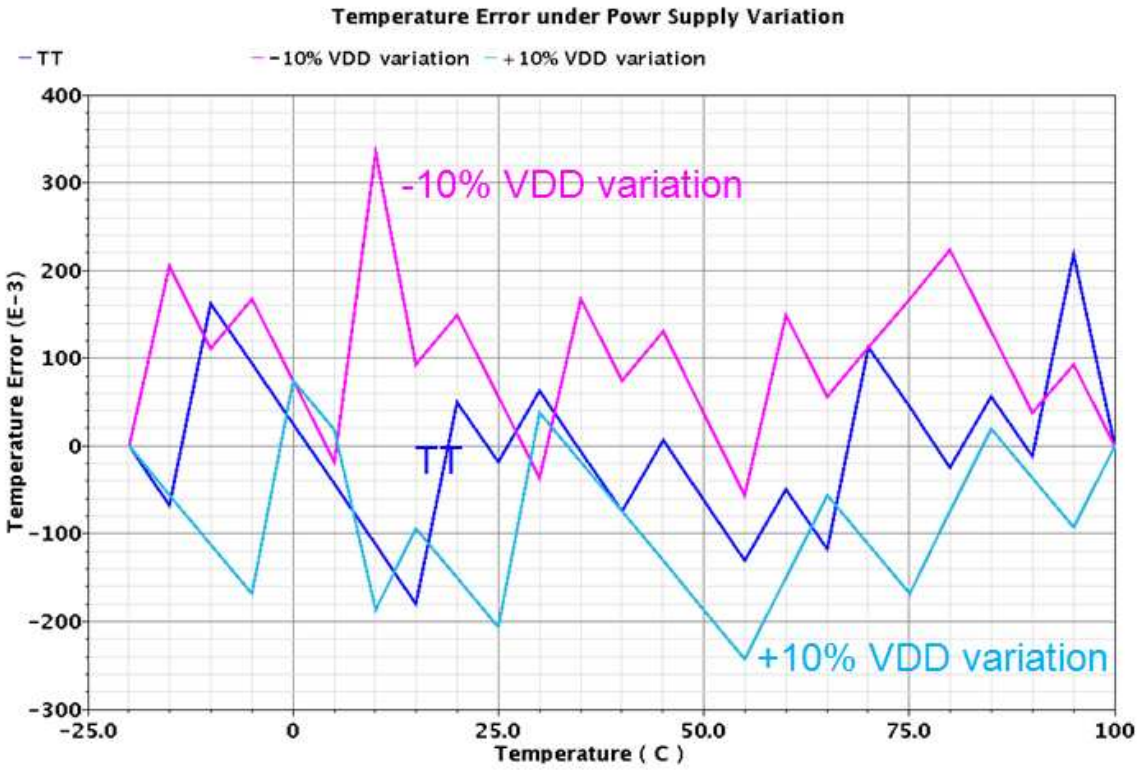

Figure 7.9 Temperature error under ±10% V_{DD} variation

A summary of the performance of this circuit as obtained from computer simulations along with the basic operating conditions are summarized in Table 7.2. The performances have been compared with the published state of art designs in temperature sensor with digital outputs in Table 7.3. Designs in [4][6][7] either utilized bandgap voltage references or accurate external reference. The proposed temperature to digital structure does not require any accurate voltage or current references, and therefore does

not have bipolar devices or parasitics BJTs, which are usually larger than standard CMOS devices. Also, the proposed design does not require any traditional ADCs, such as SARADCs and ΣΔADCs, which usually require area-consuming capacitors. In the simulation, the circuit also demonstrates robust performances under different five process corners. For continuous operation, the power consumption of the circuit is 324µW. However, a low-power mode is provided where the part is normally shut down and powers up when required to take an on-die temperature measurement. In this type of mode, assuming that conversion rate is at 10 samples/s, the estimated average power can be reduced to around 1 micro-watt range.

Table 7.2 Key circuit performances

Parameters	*Performances*
Process (µm)	0.18
V_{DD} (V)	1.8
Temperature Range (°C)	-20 ~ 100
Maximum Temp Error (°C) at typical condition	0.2
Maximum Temp Error (°C) at worst corner	0.5
Maximum Temp Error (°C) at VDD± 10% variation	0.35 @90% VDD, 0.25 @110%VDD
Total gate area of analog front end (µm^2)	1200
Output digital code range at TT w.r.t full scale range	20%
Power consumption of analog front end (µW)	324

Table 7.3 Comparison with recent integrated temperature to digital sensors

Sensors	Max Temp Error (°C)	Area (mm^2)	Temperature Range (°C)	Power Consumption	Conversion Rate (samples/s)	CMOS technology
PN junction based sensor + ΣΔADC[4]	±1	~1.3	-40~120	7 µW	50 samples/s	2 µm
PN junction based sensor +SARADC[6]	±1	3.32	-55~125	1 µW (3mW@ continuous operation)	10 samples/s	0.6 µm
PN junction based + ΣΔADC[7]	0.1 (one corner, 24 samples)	4.5	-55~125	> 187.5 µW	0.125~30 samples/s	0.7µm
TDC [9]	-0.7~+0.9	0.175	0~100	10µW	10K samples/s	0.35µm
Ring oscillator based [20]	+2.75 ~ -2.9	0.0013	-40~110	400µW	366 Ksamples/s	65nm
PN junction based[22] For PowerPC microprocessor	±4 with trim	N/A	10~100	N/A	N/A	N/A
Proposed	0.2(TT) 0.5 (five corners)	Analog front end 0.0048 Digital part (estimated) 0.006	-20~100	324µW (continuous operation)	66 Ksamples/s	0.18 µm

IV. SOME PRACTICAL CONSIDERATIONS

All results presented in this chapter were based upon either analytical formulations or computer simulations. The simulated results indicate linearity performance that exceeds that of any published results that are based upon CMOS devices. Specifically excluded are comparisons with sensor structures that sense temperature through pn junctions which can be viewed as parasitic devices in a CMOS process. Ultimately, experimental verification is essential to validate the proposed circuit. However, even if the specific circuit discussed in this chapter does not have experimental performance at the level predicted by simulations, the basic concept of designing a temperature to digital converter that does not require any voltage or current reference can be used with other temperature-sensing structures as well.

Although designers have come to expect good agreement between simulated and measured results throughout the industry, this agreement is contingent upon having good device models. In many applications, the models provided by industry are very good thus providing good agreement between simulated results and measured results. But, the literature is rather sparse on discussing the validity of existing temperature-dependent models for process parameters in CMOS technology. Existing models show extreme linearity with temperature of the threshold voltage of both n-channel and p-channel devices. But whether existing model is sufficient to predict linearity in the threshold voltage to the 0.2^{o}C or less level has not been discussed in the literature and this must be verified experimentally.

If either nonlinearities in the temperature coefficients of threshold voltage become significant or if the uncorrelation of the temperature coefficients become significant, some additional considerations may be required to obtain experimental performance at the level predicted by computer simulations but details cannot be provided until these effects are quantified. However, a batch calibration or a two-point or three-point calibration is a standard method of addressing these problems, and likely will be applicable to the proposed circuit if needed.

V. CONCLUSIONS

In this chapter, a new temperature-to-digital converter architecture called a "master slave hybrid" temperature sensor was introduced. Unlike conventional BJT based temperature sensors with digital readout, the new topology does not require any conventional constant reference voltage or reference current; it does not attempt to make any node voltage or branch current constant or precisely linear to temperature; it does not

need any conventional ADC or DAC; yet it generates a digital output code that is extremely linear to temperature over a wide temperature range and over wide process variations. Due to its low complexity and high accuracy, it provides potential for on-die temperature measurements that are for high-accurcy, small-area, and multiple-site temperature measurement.

REFERENCES

[1] A. Bakker and J. Huijsing. *High-Accuracy CMOS Smart Temperature Sensors*. Boston: Kluwer Academic, 2000.

[2] Smartec, "SMT160-30 Smart Temperature Sensor" http://www.smartec.nl, Nov.1999

[3] National Semiconductor, "LM75-Digital temperature sensor and thermal WATCHDOG with two wire inteface", http://www.national.com/ds/LM/LM75.pdf , Jan. 29th, 2009.

[4] A. Bakker and J. Huijsing, "Micropower CMOS Temperature Sensor with Digital Output", *IEEE Journal of Solid State Circuits*, Vol.31, pp. 933-937, July 1996.

[5] A. Bakker, "CMOS Smart Temperature Sensors – An Overview", *Proc. IEEE Sensors Conference*, Vol. 2, pp 1423-1427, 2002

[6] Tuthill, M., "A Switched-Current, Switched-Capacitor Temperature Sensor in 0.6um CMOS", *IEEE J. Solid-State Circuits*, Vol. 33, no.7, pp. 1117-1122, July 1998.

[7] M.A.P. Pertijs, A. Bakker, J. H.Huijsing, "A highaccuracy temperature sensor with second-order curvature correction and digital bus interface." *ISCAS*, Vol.1, pp. 368 –371, 2001.

[8] M. Pertijs, K. Makinwa, J. Huijsing, “A CMOS Smart Temperature Sensor with a 3σ Inaccuracy of ±0.1°C from -55°C to 125°C”, *IEEE Journal of Solid State Sircuits*, Vol. 40, pp. 2805-2815, Dec. 2005.

[9] P. Chen, C.-C. Chen, W.-F. Lu, and C.-C. Tsai, “A time-to-digitalconverter-based CMOS smart temperature sensor,” *IEEE J. Solid-State Circuits*, Vol. 40, no. 8, pp. 1642–1648, Aug. 2005.

[10] C.-C. Chen, P. Chen, A.-W. Liu, W.-F. Lu, and Y.-C. Chang, “An accurate CMOS delay-line-based smart temperature sensor for low-power low-cost systems,” *Meas. Sci. Technol.*, Vol. 17, no. 4, pp. 840–846, Apr. 2006.

[11] P. Chen, M. Shie, Z. Zheng, Z. Zheng, and C. Chu, “A Fully Digital Time-Domain Smart Temperature Sensor Realized With 140 FPGA Logic Elements,” *IEEE Transactions on Circuits and Sys.*, Vol. 54, No. 12, pp. 2661-2668, Dec 2007.

[12] P. Chen, T-K Chen, Y-S. Wang, and C-C. Chen, “A Time-Domain Sub-Micro Watt Temperature Sensor with Digital Setpoint Programming,” *IEEE Sensor Journal.*, Vol. 9, No.12, pp. 1639-1646, Dec 2009.

[13] Y. P. Tsividis, *Operation and Modeling of the MOS Transistor*. New York: McGraw-Hill, 1987.

[14] K. R. Laker andW. M. C. Sansen, *Design of Analog Integrated Circuits and Systems*. New York: McGraw-Hill, 1994.

[15] Filanovsky, I. M. and Allam, A., “Mutual compensation of mobility and threshold voltage temperature effects with applications in CMOS circuits,” *IEEE Trans. Circ. Syst.-I: Fundament. Theor. Appl.* 48, 7, pp.876—884, July 2001.

[16] B. Razavi and B. A. Wooley, “Design techniques for high-speed highresolution comparators,” *IEEE J. Solid-State Circuits*, Vol. 27, no. 6, pp. 1916–1926, Dec. 1992.

[17] A. J. H. Atherton, “An offset reduction technique for use with CMOS integrated comparators and amplifiers.” *IEEE J. Solid-State Circuits*, Vol.21, pp. 1168-1175, August 1992.

[18] H. Lee, “A 12-b 600 ks/s digitally self-calibrated pipelined algorithmic ADC,” *IEEE J. Solid-State Circuits*, Vol. 29, pp. 509–515, Apr. 1994.

[19] K. Woo, S. Meninger, T. Xanthopoulos, E. Crain, D. Ha, and D. Ham, "Dual-DLL-based CMOS all-digital temperature sensor for microprocessor thermal monitoring," *IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers*, pp. 68-69, Feb 2009.

[20] K. Kim, H. Lee, S. Jung, C. Kim, “366kS/s 400uW 0.0013mm2 Frequency-to-Digital Converter Based CMOS Temperature Sensor Utilizing Multiphase Clock,” *IEEE Custom Intergrated Circuits Conference*, pp. 203-206, 2009.

[21] X. Xi, M. Dunga, J. He, W.Liu and C. Hu, *BSIM4.3.0 MOSFET Model Users' Manual*. Berkeley, CA: Univ. California, pp. 12-1, 2001.

[22] H. Sanchezm, et al., “Thermal Management Systems for High Performance Power PC Microprocessors”, *Proc. IEEE Compcon*, pp. 325-330, Feb. 1997.

CHAPTER 8
CONCLUSIONS

This work has focused on the design and analyses of three fundamental building blocks: dynamic comparators, voltage references, and on-die temperature sensors.

For dynamic comparators, due to the time-varying clock signal and internal positive feedback, each transistor's operation region changes during each clock period and how it changes depends on the input signal value during that clock period. It is very difficult to predict analytically the offset in such time-varying, non-linear system. We proposed a "balanced mode" strategy to overcome the difficulty and derived explicit expressions for the offset voltage. Two types of offset were analyzed: static offset caused by process parameter mismatch dominated by carriers' mobility μ and threshold voltage V_{th}; and dynamic offset caused by parasitic capacitor mismatch and capacitive load mismatch from the imbalance in component sizes, via locations, metal routings, neighborhood coupling and loading effect, and so on. Using the balanced mode method, we first analyzed static offset in one classic comparator topology – the "Lewis-Gray" comparator. The circuit is implemented in 0.25 μm and 40-nm CMOS processes. Good agreements were achieved between results from our analytical model and those from more accurate but time-consuming Monte Carlo simulations. The analytical model also gave a good prediction to the offset in the second dynamic comparator topology proposed by L. Sumanen, et al. The analytical model also provides explicit formulae of the static offset voltage, which allow designers can use as a guidance to optimize their designs. The potential of the proposed analytic static offset model was clearly demonstrated when it

was applied to re-size "Lewis-Gray" comparator and to achieve an offset reduction of 41% while maintaining the total area as a constant. Using the balanced mode approach, we also analyzed the capacitive mismatch at different nodes in a dynamic comparator. It has been found in this work that different pairs of nodes have different vulnerabilities to capacitive mismatch. From the analytical model, we can see that if the speed requirement can be easily met, some precisely matched capacitors can be added simultaneously at the pair of nodes to reduce offset.

Motivated by the apparent gap between bandgap measured results and predicted results from existing analytical models, we developed an explicit model for the effects of non-idealities in bandgap references. Specifically, we proposed a systematic way to determine analytically the effects of the temperature dependent non-ideal component on the inflection point location, on the curvature of bandgap curve and on the magnitude of the output voltage. Two major non-ideal components, the temperature dependent gain-determining resistors and the amplifier offset voltage V_{os}, on the temperature characteristics of basic bandgap circuits were analyzed. The effectiveness of the derived model was proved by comparing with simulation results using the BSIM3v3 model in Spectre. This new approach allows the circuit designers to have a better understanding of the main limitations of the adopted voltage references architectures and to improve their high precision reference circuit designs.

For on-die temperature sensors, we developed a class of temperature sensors that express CMOS threshold voltage V_{th} at the outputs. A sizing strategy that utilizes a combined analytical and numerical approach has been introduced, which effectively compensates both second and third-order temperature non-linearity, so that the output

voltage can be highly linear with temperature. Three structures that can extract threshold voltage were designed using 0.18 μm process. The designed circuits demonstrate extremely low maximum temperature errors at 0.04^{o}C~0.15^{o}C levels at typical conditions. In the presence of both process variation and local device mismatches, the circuits all demonstrate excellent robustness. The small area and small temperature error makes this class of temperature sensors very attractive for high-precision multi-site on-chip temperature measurements. Finally, based upon the basic sensor structures, a new topology to build highly linear on-die temperature sensors with direct digital output has been proposed. This topology comprises of a master circuit using a basic threshold extraction circuit to work as a current reference, and a slave temperature sensor blended with PMOS current DAC. A comparator compares two sensing nodes voltages in the slave sensor, and drives SAR logic to control the current steering of DAC so that equal voltages are maintained at the two sensor nodes. The output digital codes from the SAR logic, which are also the inputs to the current DAC, have a very good linear relationship with temperature—with a maximum temperature error at 0.15^{o}C level. Meanwhile, this topology does not require any additional constant reference voltage or current. It does not attempt to make any node voltage or branch current constant or precisely linear to temperature. It also does not require any conventional ADC.

ACKNOWLEDGEMENTS

I have been picturing myself writing the "acknowledgement" page of my dissertation, for I owe so much gratitude to so many people in my life. They made the long process of pursuing a Ph.D. much more enjoyable and pleasant than it could have been otherwise.

First, I would like to thank my two major professors—Dr. Degang Chen and Dr. Randall Geiger, who are extraordinary mentors in my life. As the famous Chinese idiom says: "He who teaches me for one day is my father for a lifetime." I am extremely grateful for everything they taught me, especially during each weekly meeting.

I would also thank my committee members, Dr. Robert Weber, Dr. Chris Chu, Dr. Kenneth Koehler for taking their precious time to provide insightful comments and guidance on my research.

Many thanks go to my co-workers in the analog VLSI research group. During the years of my study, they collaborated with my work, accompanied me and provided suggestions and helps.

I deeply appreciate the support from my beloved family and dearest friends. Thank you for bearing with me when I complained about the stress on the work and the future. Thank you for sharing joys and sorrows. It was a blessing that I have you in my life.